





الإحصاء التطبيقي

المفاهيم الأساسية وأدوات التحليل الإحصائي الأكثر استخداماً في الدراسات والبحوث الاجتماعية والإنسانية باستخدام spss

تأليف
د. سعد بن سعيد القحطاني





بسم الله الرحمن الرحيم



مركز البحوث

الإحصاء التطبيقي

المفاهيم الأساسية وأدوات التحليل الإحصائي الأكثر استخداماً في الدراسات والبحوث الاجتماعية والإنسانية باستخدام SPSS

تأليف

د. سعد بن سعيد القحطاني

١٤٣٦هـ - ٢٠١٥م

بطاقة الفهرسة

ⓗ معهد الإدارة العامة، ١٤٣٦هـ.

فهرسة مكتبة الملك فهد الوطنية أثناء النشر.

القحطاني، سعد بن سعيد.

الإحصاء التطبيقي: المفاهيم الأساسية وأدوات التحليل الإحصائي / سعد بن سعيد القحطاني - الرياض، ١٤٣٦هـ

٣٣٢ ص؛ ١٧ × ٢٤ سم.

ردمك: ٥ - ٢٣٤ - ١٤ - ٩٩٦٠

١- الإحصاء التطبيقي ٢- التحليل الإحصائي

أ. العنوان.

ديوي: ٥١٩,٥ ١٤٣٦/٤٧٣٣

رقم الإيداع: ١٤٣٦/٤٧٣٣

ردمك: ٥ - ٢٣٤ - ١٤ - ٩٩٦٠

قائمة المحتويات

الموضوع	الصفحة

الموضوع	الصفحة

الموضوع	الصفحة

الموضوع	الصفحة

الموضوع	الصفحة

الموضوع	الصفحة
الفصل الثامن: مقاييس الارتباط بين المتغيرات Measures of Association (Correlation)	٢٠٣
مقدمة	٢٠٣
دليل إرشادي لاختيار المقياس الأنسب من بين مقاييس الارتباط بين المتغيرات	٢٠٥
تفسير قيمة معاملات الارتباط الكمية والنوعية	٢٠٩
العلاقة بين متغيرين كميين	٢١٠
الارتباط والسببية Correlation and Causality	٢١٢
أنواع العلاقات بين متغيرين X1، وY في ظل وجود متغير ثالث X2	٢١٤
مقاييس الارتباط المعلمية Parametric Correlation Measures	٢١٥
أولاً: معامل ارتباط بيرسون البسيط Pearson's Correlation Coefficient	٢١٥
كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي	٢١٦
إضاءات إحصائية حول معامل ارتباط بيرسون	٢٢٠
ثانياً: معامل الارتباط الجزئي بين المتغيرات Partial Correlation Coefficient	٢٢٢
كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي	٢٢٣
مقاييس الارتباط اللامعلمية Non-Parametric Correlation Measures	٢٣١
أولاً: معامل ارتباط سبيرمان للرتب Spearman Rank Correlation Coefficient	٢٣١
كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي	٢٣٢
إضاءات إحصائية حول معامل ارتباط سبيرمان	٢٣٥
ثانياً: اختبار مربع كاي للاستقلالية Chi Square Test for Independence	٢٣٦
إضاءة إحصائية حول مربع كاي للاستقلالية	٢٣٧

الموضوع	الصفحة

الموضوع	الصفحة

قائمة الأشكال

رقم الشكل	الشكل	الصفحة

رقم الشكل	الشكل	الصفحة

رقم الشكل	الشكل	الصفحة

رقم الشكل	الشكل	الصفحة

رقم الشكل	الشكل	الصفحة

رقم الشكل	الشكل	الصفحة

قائمة الجداول

رقم الجدول	الجدول	الصفحة

رقم الجدول	الجدول	الصفحة

المقدمة:

مما لاشــك فيه أن التحليل الإحصائي يحظى بمســتوى عال من الاهتمام من قبل الباحثين والمســتفيدين في شتى العلوم والمجالات المختلفة، وهو وسيلة هامة وأساسية في إنجاز الأبحاث والدراســات للإجابة عن تساؤلاتها وفرضياتها وهو جزء لا يتجزأ منها، فلا تكاد تخلو أي دراسة أو بحث علمي من استخدام أساليب التحليل الإحصائي المختلفة. ومن خلال ممارســات المؤلف الميدانية والخبرات العملية في مجال التحليل الإحصائي المتمثلة في تقديم الاستشارات الإحصائية ومناهج البحث وتقييم الدراسات والبحوث والتدريب والتدريس للعديد من المواد في التحليل الإحصائي نشــأت الحاجة لتأليــف هــذا الكتاب لعدة أســباب. منها ملاحظة وإدراك أن هنــاك خلطاً في فهم المصطلحات والمبادئ الإحصائية وســوءاً لاختيار أساليب التحليل الإحصائي الأنسب لموضــوع الدراســة أو البحث. كما لوحظ، مع اختيار الأســلوب الإحصائي الأنســب للتحليل، أنه ما زال هناك ســوء استخدام وتطبيق له مما يقلل من الاستفادة من نتائج التحليل وموثقيتهــا. كما لوحظ أن طريقة طرح وعرض المادة العلمية الإحصائية في كثير من المراجع الإحصائية العربية لا تخدم بشــكل فعال شــريحة كبيرة من الباحثين والمســتفيدين، لا ســيما غير الإحصائيين بالشــكل المطلوب، بل على العكس من ذلك فإنها تولد لديهم الكثير من الغموض وعدم الفهم بالشكل المرجو ومحدودية الاستفادة المطلوبة في الواقع العملي من الأســاليب الإحصائية. إن قسماً من تلك المراجع يركز علــى الجانب النظري والصيغ الرياضية أكثــر من الجانب التطبيقي مما يصعب على غيــر المتخصصين فهمها، والقســم الآخر منها يركز علــى الجانب التطبيقي أكثر من النظري ولكنها تتسم إما بالحشو الممل أو الاختصار المخل.

لذا أتى هذا الكتاب في عشــرة فصول تناولت وركزت على الأســاليب الإحصائية الوصفية والاســتدلالية الأكثر استخداماً وشيوعاً في الدراسات والبحوث الاجتماعية والإنسانية باستخدام برنامج التحليل الإحصائي SPSS وعرضها بأسلوب سهل وميسر ولغة واضحة تمكن الباحث والمستفيد من التعرف على آلية تنفيذ وإجراء تلك الأساليب وفق المنهجية العلمية الصحيحة والمتعارف عليها في التحليل الإحصائي للبيانات وكيفية تفسير نتائجها وعرضها بعيداً عن الاختصار المخل أو الإسهاب الممل.

حيــث تناول الفصــل الأول من هذا الكتــاب البحث العلمي ومفهومــه، وأهدافه، وأنواعــه، ومراحله وخطواته، وتعريف موجز ببرنامج التحليل الإحصائي SPSS. أما الفصــل الثاني فقد تناول المفاهيم الأساســية في الإحصاء. وجاء الفصل الثالث عن

العينـــات الإحصائية. أمـــا الفصل الرابع فقد اهتم بالأســـاليب الإحصائية الوصفية وتعرض الفصل الخامس لاختبارات الفرضيات الإحصائية. وفي الفصل السادس تم تناول الاختبارات المعلمية للفروق بين المتوســـطات. والفصل السابع تناول الاختبارات اللامعلمية للفروق بين المتوسطات والمتمثلة في الاختبارات البديلة للاختبارات المعلمية التي تم التطرق إليها في الفصل السادس. وأتى الفصل الثامن حول مقاييس الارتباط بـــين المتغيرات. على حين أتى الفصل التاســـع حول تحليل الانحدار الخطي. وأخيراً أتى الفصل العاشـــر ليتناول التحليل العاملي. كما أُلحق الكتاب في الملحق (١) بدليل لاختيار الاختبار الإحصائي المناســـب حســـب عدد المتغيرات المستقلة والتابعة ونوعها والهدف من الاختبار.

وقد تميـــز الكتاب عن غيره من الكتب العربية في مجـــال الإحصاء التطبيقي في عـــدة جوانب. منها، أن الأســـاليب الإحصائية الوصفية والاســـتدلالية التي تضمنها هـــذا الكتاب تم انتقاؤها بما يوفـــر احتياج البحوث والدراســـات العلمية الاجتماعية والإنســـانية من أســـاليب التحليل الإحصائي، ومن ثم طرحها بطريقة تطبيقية ميسرة ومباشـــرة يمكن للباحثين والمستفيدين استخدامها نموذجاً عملياً، وذلك وفق المنهجية التالية: الهدف من الأســـلوب الإحصائي، متى يتم اســـتخدامه، شـــروط استخدامه، طريقة تنفيذه خطوة بخطوة من خلال تطبيق عملي تفاعلي باستخدام برنامج SPSS، وأخيراً قراءة مخرجات التحليل الإحصائي وتفســـيرها وكتابة نتائج التحليل وتوثيقها بأســـلوب علمي. كما أن الكتاب اشـــتمل على العديد من «الإضاءات الإحصائية» التي تســـلط الضوء على المفاهيم والأســـاليب الإحصائية التي يساء اســـتخدامها من قبل الباحثين والطلاب والتســـاؤلات الجدلية حول بعض القضايا الإحصائية، حيث تناول وناقش تحديد الحد الأدنى لأحجام العينات العشـــوائية البسيطة أو المركبة والعينات غير العشـــوائية اللازم توفره لإجراء الدراسات والبحوث التي تحتوي على العديد من المتغيرات الكمية والنوعية المختلفة. كما تم تســـليط الضوء على اختبارات الفرضيات في حالة بيانات الحصر الشامل وأنسب اختبارات المقارنات البعدية في تحليل التباين، وغيرها من التســـاؤلات والمفاهيـــم الأخرى التي تم مناقشـــتها وتوضيح الرأي حولها وتضمينهـــا في كتاب واحد توفر على الباحث عناء التنقل بين الكتب الأجنبية والعربية للحصول على استيضاح حول تلك القضايا والمفاهيم.

الفصل الأول
البحث العلمي Scientific Research

مفهوم البحث العلمي:

إن البحث هو وسيلة وليس غاية في حد ذاته فمن خلاله يسعى الباحث إلى إشباع حاجاته ورغباته من العلوم والمعرفة والتعرف على الظواهر المحيطة به من خلال وصفها ودراسة العلاقات بينها ومعرفة الأسباب والدوافع في حدوثها. إن المطلع على أدبيات البحث العلمي يجد أن هناك العديد من التعاريف المتفاوته نسبياً لمفهوم البحث العلمي. ومن هذه التعاريف ما أورده العساف والوادي (٢٠١١) بأن فان دالين عرف البحث العلمي على أنه «محاولة دقيقة ومنظمة وناقدة للتوصل إلى حلول لمختلف المشاكل التي تواجهها الإنسانية». وذكر النجار، النجار والزعبي (٢٠١٠، ٢٢) بأن البحث العلمي هو «طريقة منظمة أو فحص استفساري منظم لاكتشاف حقائق جديدة أو التثبت والتحقق من حقائق قديمة والعلاقات التي تربط بينها أو القوانين التي تحكمها بما يسهم في نمو المعرفة الإنسانية».

وعلى الرغم من تعدد تعريفات البحث العلمي إلا أنها تتفق على أن البحث العلمي (العساف والوادي، ٢٠١١، ٢٧):

١- عبارة عن محاولة منظمة تعتمد على الأساليب العلمية.

٢- يؤدي إلى زيادة الحقائق التي يعرفها الإنسان.

٣- يختبر المعارف والعلاقات والظواهر.

٤- شامل لجميع الميادين والمجالات.

أهداف البحث العلمي:

تسعى الدراسات والأبحاث العلمية إلى دراسة قضايا أو ظواهر معينة وفق طرق وخطوات علمية محددة موضوعية لا تخضع للاجتهادات الشخصية أو الارتجالية وذلك بهدف وصفها، وتحليلها، وتفسيرها، واستقراء مستقبلها في ظل المتغيرات ذات العلاقة بالظاهرة. ومن المعروف لدى الباحثين والمختصين بأن أهداف البحث العلمي تتمثل في النقاط الأربع التالية (النجار وآخرون، ٢٠١٠، ٢٥):

١- الوصف Description: ويقصد به وصف الظاهرة أو القضية محل الدراسة من خلال جمع البيانات اللازمة حولها وتجميعها وتبويبها لتحديد ملامح وخصائص تلك الظاهرة أو القضية. وكمثال على ذلك وصف خصائص المسجلين في برنامج حافز خلال فترة زمنية محددة.

٢- التفسير Explanation: ويقصد به اكتشاف وتحديد العوامل أو الأسباب المؤدية إلى حدوث ظاهرة أو قضية ما والعمل على تحليلها والربط بينها بهدف الوصول إلى تفسير واضح وعلمي حيال تلك الظاهرة أو القضية. وكمثال على ذلك دراسة أسباب ودوافع الفساد الإداري والمالي في الأجهزة الحكومية في بلد ما.

٣- التنبؤ Prediction: من أهداف البحث العلمي التنبؤ واستقراء المستقبل لظاهرة أو قضية ما. وتعتبر عملية التنبؤ بالقيم المستقبلية من الأهداف الهامة للبحث العلمي فهي تدعم عملية القرارات المستقبلية. ودرجة دقة التنبؤ تعتمد على عدة عوامل منها الدقة في تحديد نماذج التنبؤ وثبات الظاهرة محل الدراسة عبر الزمن. ومن أمثلة التنبؤ: التنبؤ بأعداد العاطلين عن العمل، أعداد الخريجين من جامعة ما، سعر سهم ما ... إلخ.

٤- الضبط أو التحكم Control: وهو الهدف النهائي للعلم، فمتى وصل الباحث إلى هذه المرحلة فإنه يعتبر قد امتلك السيطرة على عوامل وأسباب الظاهرة أو القضية محل الدراسة وأصبح بإمكانه التحكم في مدخلاتها ومخرجاتها كما يريد. وكمثال على ذلك التحكم في أوزان الأشخاص من خلال برامج غذائية ورياضية معينة.

أنواع البحوث العلمية:

سيجد المطلع على أدبيات مناهج البحث العلمي أن هناك تصنيفات عدة لها، إلا أن بعض تلك التقسيمات والتصنيفات لاقت قبولاً وشبه اتفاق أكثر من غيرها بين المختصين في مجال مناهج البحث العلمي. وهذه التصنيفات تعتمد على معايير مختلفة مثل طبيعة البيانات، طبيعة البحث، الغرض من البحث، الزمن، المكان، منهج البحث، مستوى البحث (قنديلجي والسامرائي، ٢٠٠٩). وبناء على تلك المعايير فإنه يمكن تقسيم البحوث العلمية حسب معيار طبيعة البيانات إلى قسمين رئيسيين (انظر الشكل ١-١) هما: البحوث الكمية والبحوث الكيفية، ويتفرع من كل قسم تقسيمات فرعية أخرى تعتمد في تقسيمها على المعايير الأخرى السابق ذكرها.

أولاً - البحوث الكمية Quantitative Research:

هي البحوث التي تسعى لوصف الظواهر والعلاقات بينها وتفسيرها للوصول إلى نتائج معينة باستخدام الأساليب الكمية. فالبحوث الكمية تتخذ من الأرقام والأساليب الكمية أساساً لها وتستخدم في الغالب البيانات الكمية (الرقمية) التي يتم الحصول عليها من وحدات الدراسة ومن ثم استخدامها في تحليل وتفسير الظاهرة محل الدراسة. وتعتمد البحوث الكمية على مبدأ الاستنباط deduction الذي يبدأ بالنظريات لاستنباط الفرضيات منها، ومن ثم العمل على اختبار تلك الفرضيات من خلال جمع البيانات وتحليلها. ويتم عادة استخدام العينات العشوائية (الاحتمالية) وأدوات جمع البيانات المقننة في البحوث الكمية وتشمل البحوث الكمية الأنواع التالية:

١- البحوث المسحية Survey Research: وهي من أكثر مناهج البحث الكمي استخداماً، وفيها يتم جمع البيانات والمعلومات من عناصر مجتمع أو عينة الدراسة لتحليلها واستخدامها في الإجابة عن أسئلة وفرضيات الدراسة. ولهذا النوع من البحوث عدة أنواع هي:

١- البحوث المسحية الوصفية: ويهتم بوصف الظاهرة محل الدراسة وتحديد خصائصها والعلاقات بين أبعادها والعوامل المؤثرة فيها.

٢- البحوث المسحية الارتباطية: ويهتم بدراسة العلاقات الارتباطية بين متغيرات الظاهرة محل الدراسة.

٣- البحوث المسحية التنبؤية: وتهدف إلى الخروج بتنبؤات مستقبلية بناء على البيانات والمعلومات في الواقع الراهن.

٤- البحوث المسحية التطورية: وتعمل على دراسة مراحل تطور وتغير الظاهرة خلال فترات زمنية متتالية يحددها الباحث.

٢- البحوث التجريبية Experimental Research: وهي طريقة علمية منظمة للبحث تعتمد أسلوب التجربة، وفيها يعمل الباحث على التلاعب بمتغير أو أكثر من متغيرات الدراسة ليتحكم ويقيس التغيرات التي تطرأ على متغير أو متغيرات أخرى في الدراسة لاختبار فرضيات الدراسة. والبحوث التجريبية تنقسم إلى بحوث تجريبية فعلية True Experimental Research، وبحوث شبه تجريبية Quasi-Experimental Research. وتختلف عن بعضها بشكل أساسي في درجة

التحكم بمتغيرات الدراسة والتوزيع العشوائي لوحدات التحليل أو المعاينة على المجموعات في الدراسة.

٣- البحث المقارن Comparative Research: ويهدف إلى المقارنة بين ظاهرتين أو أكثر من خلال التعرف على أوجه الشبه والاختلاف بينها، فقد تكون المقارنة لظاهرة ما في مجتمعين مختلفين أو مقارنة ظاهرة في نفس المجتمع ولكن في فترات زمنية مختلفة. ومثال على ذلك مقارنة الأداء بين جهازين حكوميين أو أكثر في بلد معين.

ثانياً - البحوث الكيفية (النوعية) Qualitative Research:

هي تلك البحوث التي تركز على دراسة الظواهر والأفراد والمجتمعات والمؤسسات في مختلف مجالات العلوم ووصفها والفهم الأعمق لها، وهي عادة لا تركز على التعميم بل تركز على دراسة أعمق للحالات وتوسيع نتائجها حتى يمكن الاستفادة منها في حالات وظروف أخرى مماثلة (قنديلجي، السامرائي، ٢٠٠٩). والبحوث الكيفية في الغالب تعتمد على بيانات نوعية تكون على شكل ملاحظات وتعليقات وآراء مكتوبة أو مشاهدة أو مسموعة يتم الحصول عليها من وحدات الدراسة ومن ثم استخدامها في تحليل وتفسير الظاهرة محل الدراسة. وتعتمد البحوث الكيفية على مبدأ الاستقراء Induction الذي يتم فيه ملاحظة الظواهر الملموسة ومن ثم بناء فرضيات حول العلاقة بين تلك الظواهر ومتغيراتها ومن ثم جمع البيانات حولها وتحليلها لاختبار صحة تلك الفرضيات للتوصل إلى تعميم. وللبحوث الكيفية أنواع منها:

- بحوث دراسة الحالة Case Study: وهي تركز على إعطاء تفاصيل دقيقة عن حالة أو أكثر. فمثلاً يمكن إجراء دراسة حالة على منظمة طبقت نظام الحكومة الإلكترونية.
- البحوث الإثنوجرافية (دراسة الأعراق والأجناس) Ethnography Research: وهي ذلك النوع من البحوث النوعية التي تركز على وصف ثقافة مجموعة من الأشخاص من جميع الجوانب المتعلقة بالعادات والتقاليد والقيم والاتجاهات والممارسات واللغة والنواحي المادية.
- منهج دراسة الظواهر Phenomenology: وهذا المنهج يركز على دراسة ووصف الظواهر التي ندرك وجودها حولنا في العالم الذي نعيش فيه، ولكنها غير مفهومة لنا بشكل كامل مثل الأحداث، المواقف، الخبرات أو المفاهيم. وكمثال على ذلك ظاهرة التعصب القبلي، المخدرات، الخوف من مادة الرياضيات، ... إلخ.

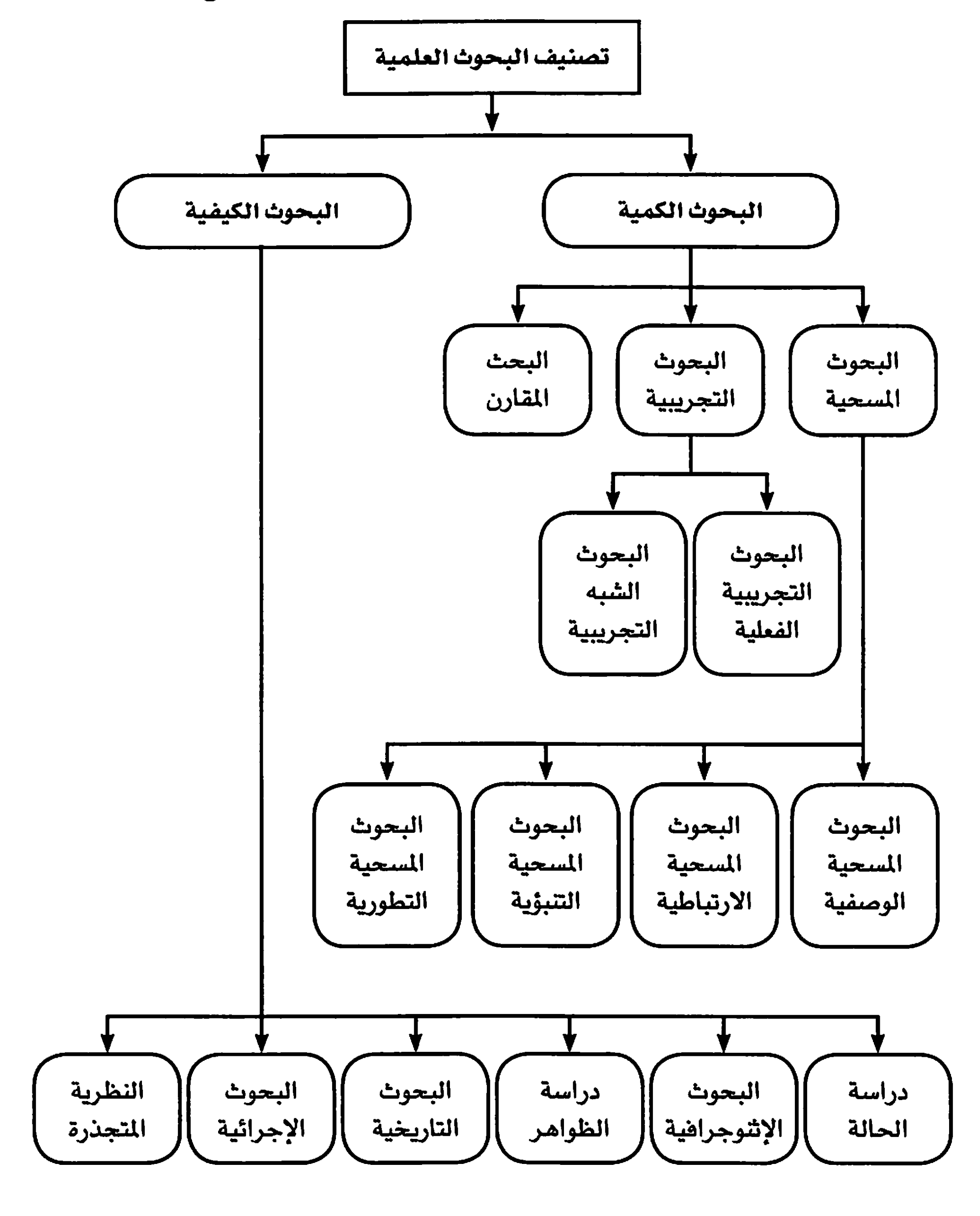
- البحوث التاريخية أو الوثائقية Historical Research: وهي تعنى بدراسة الأحداث التي ظهرت في الماضي.
- البحوث الإجرائية Action Research: هي البحوث التي تهدف بشكل أساسي إلى معالجة مشكلات عملية معينة في الحياة اليومية وتقديم الحلول حولها بشكل عاجل بصرف النظر عن إمكانية تعميمها.
- النظريـــة المتجذرة Grounded Theory: وهي تلــك البحوث التي تهدف إلى توليد وتطويــر النظريات من البيانات والمعلومات التي يقوم بجمعها الباحث. فمثلاً يمكن لباحـــث أن يجمع بيانات مفصلة حول «ظاهرة التســـرب الوظيفي من القطاع العام إلى الخاص» ومن ثم بناء نظرية تفســـر كيف ولماذا تحدث هذه الظاهرة ومن ثم يتم التوصل إلى نظرية «التسرب الوظيفي من القطاع العام إلى الخاص».

إن الاختيار من بين مناهج أو نوع البحوث العلمية يعتمد على عدة نقاط منها:

- الفترة الزمنية التي يتناولها موضوع البحث، الماضي أم الحاضر أم المستقبل.
- المكان الذي سيتم فيه البحث. هل البحث مكتبي بحت أم ميداني؟
- الغرض من البحث. هل الهدف منه الوصف فقط أم الوصف والتحليل والتفسير والتنبؤ؟
- التكلفة وأهمية البحث.
- مستوى المعرفة والمعلومات المسبقة عن موضوع البحث.

شكل رقم (١-١)

تصنيف البحوث العلمية (الشكل من إعداد المؤلف استناداً إلى المراجع الواردة أعلاه)



مراحل وخطوات البحث العلمي:

غالبـاً يمر البحث العملي بعدة مراحل متتالية (كما يوضحه شـكل ١-٢) مرتبطة ومكملة بعضها بعضاً، وهي:

١- مرحلة الشعور بمشكلة الدراسة وتحديدها:

ففي هذه المرحلة على الباحث أن يحدد مشـكلة الدراسـة بعد الشعور بها تحديداً واضحاً يمكّنه من وضع الأهداف والأسئلة والافتراضات المتعلقة بها.

إن مصادر مشكلة الدراسة البحثية متعددة منها:

- الدراسات السابقة أو الندوات والمؤتمرات العلمية.
- الكتب والمراجع والتقارير الإحصائية.
- استشارة الخبراء والمختصين الذين لديهم معرفة علمية أو عملية في مجال ما.
- الخبرة العملية والميدانية، ومثال على ذلك مشكلة «أسباب التسرب الوظيفي».
- وسائل الإعلام والإنترنت.

وقد ذكـر عليان وغنيم (٢٠٠٨) بعـض الصفات التي ينبغي توافرها في مشـكلة الدراسة كي يتمكن الباحث من دراستها، ومنها:

- أن تكون المشكلة حديثة وأن تمثل إضافة علمية مفيدة لحقل المعرفة.
- أن تكـون الدراسـة ضمن إمكانيات الباحث المتعلقة بالدعـم المادي، الوقت، الجهد المبذول، ومدى توفر البيانات والمعلومات اللازمة للدراسة.
- أن يكون لدى الباحث اهتمام بمشكلة الدراسة وقدرة على تنفيذها.
- عدم الإضرار بمصلحة خاصة أو عامة كالمصلحة المالية، الجسدية، أوالصحية وغير ذلك مما له أثر سلبي على الفرد أو المجتمع.
- أن لا تتعارض مشـكلة الدراسة مع المبادئ السياسية أو الشرعية أو القانونية، فمثلاً تعتبر دراسة «أثر حرية خروج الموظفين أثناء الدوام على مستوى الإنتاجية في القطاع الحكومي» غير مناسبة لتعارضها مع نظام الدوام الرسمي في القطاع الحكومي.

وعادة ما يتم طرح مشكلة الدراسة في هيئة سؤال عام واضح ومحدد مثل: ما هو أثر تطبيق منهجية قياس الأداء على أداء الأجهزة الحكومية في المملكة العربية السعودية؟

٢- تحديد أهداف الدراسة وصياغة أسئلتها وفرضياتها:

بعد مرحلة تحديد وصياغة المشكلة البحثية يتم الانتقال إلى المرحلة التي تليها وهي مرحلة تحديد أهداف الدراسة وصياغة أسئلتها وفرضياتها. هذه المرحلة تتطلب من الباحث أن يراجع مراجعة مكثفة وكافية الأدبيات المتعلقة بموضوع بحثه للوقوف على آخر ما تم بحثه في مجال موضوعه وما تم التوصل إليه، ومن ثم ربط ذلك بموضوع بحثه والاستفادة منه في صياغة أسئلة وفرضيات دراسته في إطار هدف أو أهداف دراسته. إن أسئلة الدراسة وفرضياتها تنشأ من أهداف الدراسة لذلك لابد من تحديد أهداف الدراسة تحديداً واضحاً. فمثلاً من الأهداف الممكنة لدراسة ميدانية بعنوان «الرضا العام للمستفيدين من خدمات الأحوال المدنية بمدينة الرياض» ما يلي:

- قياس مستوى رضا المستفيدين من خدمات الأحوال المدنية بمدينة الرياض.
- تحديد العوامل المؤثرة في رضا المستفيدين من خدمات الأحوال المدنية بمدينة الرياض.
- ولتحقيق أهداف الدراسة فإنه يتعين على الباحث أن يحدد أبعاد مشكلة الدراسة وأن يترجم أهدافها من خلال وضع أسئلة الدراسة. إن أسئلة الدراسة الجيدة تمتاز بعدة خصائص منها:
  - الوضوح والدقة في الصياغة والعلاقة المباشرة بالدراسة.
  - قابلة للدراسة والاختبار وتترجم أهداف الدراسة إجرائياً.
  - تساعد الباحث على اختيار الطريقة والأداة المناسبة لجمع البيانات.
  - تمكن الباحث من وضع تصور مبدئي للأدوات الإحصائية التي سيتم استخدامها في الإجابة عن أسئلة الدراسة.

٣- جمع البيانات:

تعتبر هذه المرحلة من أطول مراحل البحوث العلمية وأكثرها تكلفة. وتتضمن هذه المرحلة عدة جوانب تتمثل في تصميم الدراسة بالشكل الصحيح والمناسب وتحديد منهج الدراسة، تحديد مجتمع الدراسة، تحديد أسلوب المعاينة المناسب، أسلوب جمع البيانات، تصميم أداة جمع البيانات، تحديد مصادر البيانات وتحديد متغيرات الدراسة ومستوى قياسها واختيار مفردات العينة.

٤- تحليل البيانات وتفسيرها:

وفــي هذه المرحلة يتم اســتخدام أســاليب التحليــل الإحصائية ســواء الوصفية Descriptive منها أو الاســتدلالية Inferential أو كليهما للإجابة عن أسئلة الدراسة واختبار فرضياتها.

٥- مناقشة النتائج والتوصيات Results Discussion and Recommendations:

وهــي تعتبر المرحلــة الأخيرة في البحــوث العلمية، وجرت العادة على اســتهلال هذه الجزئية من البحث بإعطاء نبذة موجزة عن مشــكلة الدراسة وعن بعض النقاط الأساسية في إجراءاتها. ويمكن تجزئتها إلى جزأين هما: النتائج والتوصيات.

الجزء الأول - مناقشة النتائج:

وفــي هذا الجزء يعمل الباحث على مناقشــة وكتابة نتائج دراســته والخروج منها بالاســتنتاجات المســتخلصة من تلك النتائج ذات الصلة بأهداف الدراسة، وذلك بعد الانتهاء من عمليات التحليل الإحصائي المتعلقة بأســئلة وفرضيات الدراســة. ومما لا شك فيه أن الباحث في هذه المرحلة سيكون لديه إلمام تام بموضوع دراسته وأهدافها وفرضياتها واطلاع واســع على الدراســات السابقة ذات العلاقة بدراسته. والدراسة البحثية الجيدة ينبغي أن تتوفر فيها النقاط التالية:

١- كتابة النتائج حسب الأهمية والتسلسل المنطقي لها ومن حيث أهمية وقوة ارتباطها بمشكلة الدراسة. كما ينبغي التركيز والابتعاد عن التفصيل الممل في كتابة النتائج والذي قد يفقد القارئ فهم واســتيعاب نتائج الدراسة، ومن ثم يقلل من الاستفادة المرجوة من نتائج تلك الدراسة.

٢- تشخيص وتفسير النتائج التي توصل إليها الباحث بشكل واضح ومفيد وفي إطار مشكلة الدراسة. فينبغي على الباحث التركيز على تفسير النتائج ضمن نطاق موضوع الدراسة، أهدافها، أسئلتها وفرضياتها حتى لا تخرج الدراسة عن مجال إطارها العام.

٣- ربط نتائج الدراســة بالدراسات السابقة ذات الصلة، حيث ينبغي على الباحث أن يوضح أثناء كتابته لنتائج دراســته العلاقة وأوجه الاتفاق والاختلاف بينها وبين ما توصلت إليها الدراسات والأبحاث الأخرى المتعلقة بموضوع بحثه.

٤- الموضوعيــة والمصداقيــة في طرح النتائج ســواء كانت ســلبية أو إيجابية وتجنب أسلوب التحيز والمجاملات.

٥- الإشـــارة إلى حدود الدراســـة البحثيـــة Reasearch Study Limitations. وهنا ينبغي على الباحث أن يوضح حدود دراســـته كالحدود المكانية، الزمانية، الثقافية وما إلى ذلك حسب طبيعة وإمكانيات دراسته المادية وغير المادية والزمنية وحسب طبيعة مجتمع الدراســـة وخصائصه. فلكل دراســـة محدداتها الخاصة والمختلفة عن غيرها من الدراسات الأخرى. وتكمن أهمية توضيح حدود الدراسة في مدى الاستفادة من نتائجها وتحديد نطاق تعميمها.

الجزء الثاني - تقديم التوصيات:

وعادة يأتي تقديم التوصيات والاقتراحات بعد كتابة نتائج البحث. وعند كتابة هذه الجزئية من البحث ينبغي مراعاة النقاط التالية (القحطاني، ١٤٣٢):

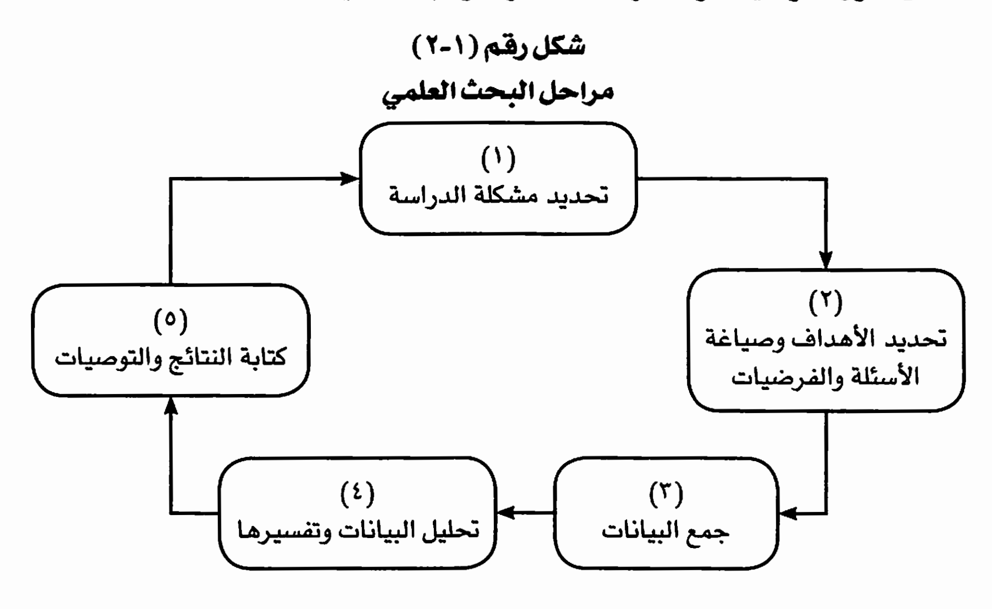
١- تقديم التوصيات والاقتراحات بناء على النتائج والاستنتاجات التي رصدها الباحث، وما لديه من خبرة في موضوع دراسته.

٢- تجنب أسلوب العموميات في التوصيات والاقتراحات.

٣- أن تكون التوصيات والاقتراحات معقولة وقابلة للتنفيذ.

شكل رقم (١-٢)

مراحل البحث العلمي



أدوات جمع البيانات Data Collection Tools:

تعــرف أداة جمع البيانات على أنها وســيلة للحصول على البيانات أوالقياســات والمعلومــات المطلوبة لدراســة ظاهرة أو قضيــة معينة. ولكل علم مــن العلوم أدواته الخاصة في الحصول على البيانات، فمثلاً في المجال الطبي يتم اســتخدام الثرمومتر لقياس درجة حرارة الأشــخاص، وفي المجال الهندســي يتم استخدام المسطرة المترية لقياس الأطوال والأبعاد، وفي المجال التعليمي يتم استخدام اختبارات التحصيل لقياس المستوى المعرفي أو التحصيلي للتلاميذ، وفي مجال الدراسات الاجتماعية أو الإدارية وغيرها من المجالات الأخرى يتم استخدام الاستبانات لقياس ظاهرة معينة.

وهناك مجموعة من أدوات جمع البيانات يمكن استخدامها في أي مجال من مجالات العلوم المختلفة منها (Johnson, Christensen, 2012) و(العساف، الوادي،٢٠١١).

- الاختبارات Tests: مثل الاختبــارات المعيارية الخاصة وكمثال على ذلك اختبارات الذكاء، المهارات، القدرات، الاختبارات التحصيلية، ... إلخ.
- الاستبانة Questionnaire.
- المقابلة Interview.
- الملاحظة Observation.
- الوثائق Documents.

ولــكل من هذه الأدوات إيجابياتها وســلبياتها ويتم الاختيــار من بين تلك الأدوات حسب المعايير التالية:

- طبيعة مشكلة البحث.
- منهج البحث المتبع.
- طبيعة مجتمع البحث وعينته وحجم العينة ونوعها.
- مدى معرفة الباحث بأداة البحث واستخدامها.
- الوقت المطلوب لجمع البيانات والجهد والتكلفة المالية.

وتعتبر الاستبانة أشــهر تلك الأدوات وأكثرها استخداماً، ولا سيما في الدراسات والبحــوث الاجتماعية والإنســانية. وتجدر الإشــارة إلى أنه قد يحتــاج الباحث إلى استخدام أكثر من أداة لجمع البيانات في بحثه.

الاستبانة Questionnaire:

هي أداة من أدوات جمع البيانات، وهي عبارة عن مجموعة من العبارات والأسئلة حول موضوع الدراسة البحثية التي تهدف إلى الحصول على بيانات معينة من المبحوثين (المستجيبين) من خلال إجاباتهم على تلك الأسئلة والعبارات، والملحق (٢) يتضمن مثالاً لاستبانة تم تطويرها واستخدامها في بحث ميداني حول «أسباب عدم الاستجابة في المسوحات البحثية» (إسماعيل، ٢٠١١). إن تصميم الاستبانة الإحصائية ينبغي أن يتم بعناية كبيرة وبطريقة مناسبة لهدف ومجتمع الدراسة، وذلك كون إجابات المستجيبين تتأثر بصياغة الأسئلة وبنوعية الأسئلة وغيرها من الاعتبارات التي ينبغي أخذها بالاعتبار أثناء مرحلة تصميم الاستبانة (فهمي، ٢٠٠٥).

قواعد صياغة الاستبانة:

لأن الاستبانة هي أداة لجمع البيانات ولأن دقة البيانات التي يتم الحصول عليها تعتبر هي غاية في الأهمية في البحث العلمي وتنعكس بشكل مباشر على دقة النتائج ومصداقيها؛ فإنه لابد للباحث من مراعاة قواعد بناء الاستبانة وأخذها بعين الاعتبار في مرحلة التصميم. والشكل (١-٣) التالي يوضح باختصار القواعد من الناحية الشكلية والمحتوى التي ينبغي مراعاتها عند تصميم الاستبانة (العساف، الوادي، ٢٠١١؛ النجار وآخرون، ٢٠١٠؛ عاروري، ٢٠١٣).

شكل رقم (١-٣)
القواعد العامة لتصميم وبناء الاستبانة

قواعد تصميم الاستبانة

الجوانب الشكلية

- احتواء الاستبانة على عنوان البحث
- ترقيم أسئلة وصفحات الاستبانة
- احتواء الاستبانة على أهداف البحث
- ينبغي مراعاة ترتيب الأسئلة. الانتقال من العام إلى الخاص ومن السهل إلى الأصعب
- فقرة تشير إلى أهمية إجابات المبحوث وتقديم الشكر لهم على تعاونهم
- أن تكون الاستبانة قصيرة قدر الإمكان مع مراعاة عدم الإخلال بمحتواها
- فقرة تشير إلى سرية إجابات المبحوث واستخدامها للأغراض العلمية فقط
- الابتعاد قدر الإمكان عن الأسئلة المتفرعة
- اسم الباحث والجهة التي ينتمي إليها
- سلامة اللغة وجودة الطباعة ونوع الورق وشكله وحجمه

جوانب المحتوى (صياغة الأسئلة)

- أن تكون العبارات واضحة تناسب المستوى التعليمي للمبحوث وألا تتضمن عمليات حسابية معقدة أو تحتاج إلى تفكير عميق
- الابتعاد عن الأسئلة أو العبارات التي تحمل كلمات فضفاضة دون توضيح المفهوم الإجرائي لها مثل الرضا
- الابتعاد من صياغة الأسئلة بعبارات إيجابية أو سلبية
- تجنب الأسئلة أو العبارات التي قد تشعر المبحوث بالحرج
- عدم استخدام الأسئلة أو العبارات المركبة، أي التي تحمل أكثر من فكرة واحدة
- الابتعاد قدر الإمكان عن استخدام العبارات المنفية لأنها قد تؤدي بالمبحوث إلى الفهم بطريقة عكسية
- الابتعاد عن الأسئلة أو العبارات الإيحائية التي توجه المبحوث لإجابات معينة حسب رغبة الباحث
- الابتعاد عن استخدام الكلمات أو المصطلحات الفنية المتخصصة غير الشائعة الاستخدام
- الابتعاد عن الأسئلة أو العبارات الطويلة لأن ذلك قد يؤدي إلى إرباك المبحوث وعدم فهم المقصود منها
- عدم استخدام الأسئلة أو العبارات المعروف إجاباتها سلفاً

صدق وثبات المقياس (أداة جمع البيانات) Validity and Reliability:

مـــن الصفـــات الهامة التي ينبغي توافرهـــا في المقاييـــس أو أدوات جمع البيانات الصدق والثبات وذلك لما لها من تأثير بشـــكل مباشـــر على مصداقية نتائج الدراسة. فإذا فقد المقياس هاتين الخاصيتين فإن الدراسة المعتمدة عليها لا تعدو كونها مضيعة للوقت وعديمة الجدوى.

- **الصـــدق Validity**: ويقصـــد به إلى أي مدى يقيس المقياس أو الأداة ما صممت من أجله. فإذ صممت الاستبانة مثلاً «لقياس رضا عملاء شركة موبايلي حول الخدمات التي تقدمها» فإنها يجب أن تصمم لقياس رضا عملاء تلك الشـــركة حول الخدمات المقدمة وليس لمعرفة رأيهم في سياســـات تلك الشـــركة مثلاً أو لقياس رضا عملاء شركة الاتصالات السعودية أو زين. ولصدق المقياس عدة أنواع كما يوضحه الشكل (١-٤) وهي كما يلي:

- **الصدق الظاهري Face Validity**: ويسمى أيضاً صدق المحكمين ويقصد به إلى أي مـــدى يبـــدو ظاهرياً أن المقياس يقيس ما صمم مـــن أجله (Gay&Airasian, 2000). ويتـــم قياس الصـــدق الظاهري للمقياس من خلال عرضه على أصحاب الاختصاص والخبرة في مجال الدراسة أو البحث ليحكموا ظاهرياً على مصداقية المقياس ودرجة قياسه لما صمم من أجله.

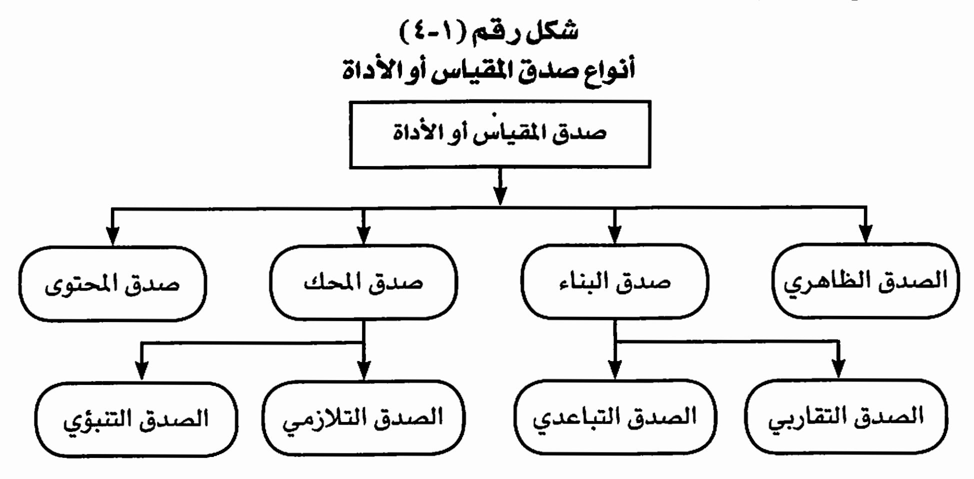
- **صـــدق المحتـــوى Content Validity**: ويقصـــد به إلى أي مدى يشـــمل المقياس خصائص الشـــيء المراد قياســـه من كل الجوانب اللازمـــة. ولأن الأداة المصممة لا تحتوي كل الأســـئلة الممكنة حول موضع البحث وإنما يكتفي الباحث بمجموعة أو عينة من الأســـئلة حول موضوع البحث، فإنه ينبغي التأكد من أن تلك المجموعة من الأسئلة كافية لتغطية محتوى موضوع البحث. فمثلاً في المثال السابق «قياس رضا عملاء شركة موبايلي حول الخدمات التي تقدمها» ينبغي أن تغطي الاستبانة جميـــع جوانب الخدمات المقدمة للعميـــل فلا تقيس الاســـتبانة خدمة بيع المنتج وتســـقط خدمة ما بعد بيع المنتج. ويتم قياس صدق المحتوى للمقياس من خلال عرضـــه على أصحاب الاختصاص والخبرة في مجال الدراســـة أو البحث للتأكد من محتواه.

- **صـــدق المحـــك Criterion-Related Validity**: ويعرفه Huck (2012) على أنه الاســـتدلال على صدق أداة جديدة مجهولة الصدق بمقارنتها بأداة أخرى معلومة الصدق. ويشمل هذا النوع من الصدق: الصدق التلازمي concurrent validity، الصدق التنبؤي Predictive validity.

- صدق البناء **Construct Validity**: وهو درجة الارتباط بين نتائج المقياس أو الأداة وبين المفهوم البنائي لها (Huck, 2012). وقد أوضح Huck ثلاث طرق لقياس الصدق البنائي وسأكتفي هنا بذكر أشهرها وأكثرها استخداماً، وهي طريقة التحليل العاملي Factor Analysis والذي سيتم مناقشته بالتفصيل في الفصل العاشر ويشمل هذا النوع من الصدق: الصدق التقاربي Convergent Validity، والصدق التباعدي أو التمييزي Divergent (Discriminant) Validity.

وتجدر الإشارة إلى أن الصدق الظاهري وصدق المحتوى وصدق البناء من أكثر أنواع الصدق المستخدمة في تقييم صدق أدوات أو مقاييس جمع البيانات في الدراسات البحثية. والأنواع الأخرى يمكن ملاحظة استخدامها في مقاييس الاختبارات والمقاييس النفسية والشخصية.

شكل رقم (١-٤)
أنواع صدق المقياس أو الأداة



- الثبات **Reliability**: ويقصد به إلى أي درجة يعطي المقياس قراءات متقاربة عند تطبيقه في كل مرة Huck (2012). وسنتعرف في الفصل العاشر على المزيد حول مفهوم الثبات وكيفية حسابه باستخدام SPSS ومن ثم تفسير نتائجه وكيفية توثيقها بأسلوب علمي.

العلاقة بين الصدق والثبات:

ينبغي معرفة أن ثبات المقياس ضروري ولكنه ليس كافياً لأن يكون صادقاً. أي أنه إذا كان المقياس ثابتاً فقد يكون غير صادق، وفي المقابل عندما يكون المقياس صادقاً فيجب أن يكون ثابتاً. والشكل (١-٥) التوضيحي يصور العلاقة بين الصدق والثبات

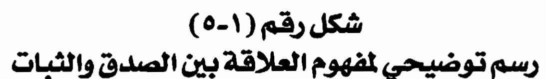
بطريقة سهلة وميسرة. فالدوائر المتداخلة تمثل درجة أو مدى صدق الاستبانة. فعند مركز تلك الدوائر تكون درجة الصدق عالية جداً وتقل درجة الصدق كلما ابتعدنا عن المركز. أما النقاط داخل تلك الدوائر فتدل على عدد مرات تطبيق أو استخدام المقياس، ومن ثم يمكن ملاحظة ما يلي:

١- الشكل (أ) يصور الحالة التي يكون عندها المقياس صادقاً وثابتاً. فكون النقاط تقع عند المركز يعني أن المقياس صادق ويقيس ما صمم من أجله (أصاب الهدف عند المركز). أما تجمع النقاط مع بعضها فهو يعني أن المقياس ثابت حيث يعطي قراءات متقاربة في كل مرة يطبق فيها.

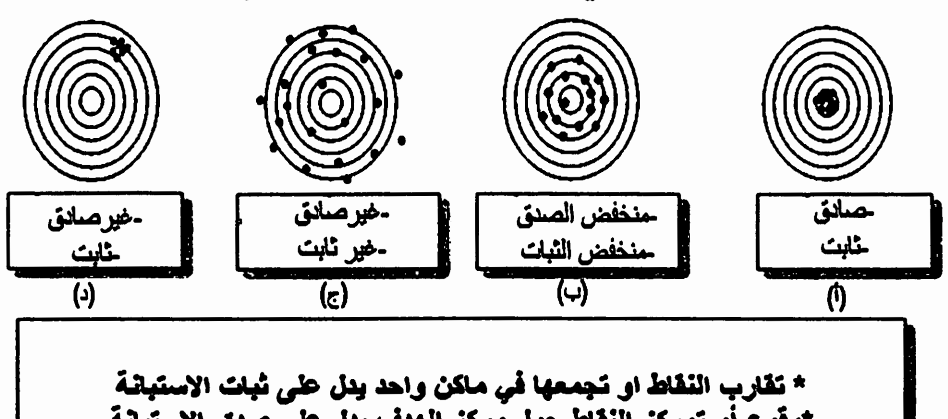
٢- الشكل (ب) يصور الحالة التي يكون عندها المقياس منخفض الصدق والثبات، لأن معظم النقاط غير متمركزة حول المركز وغير متجمعة بالقدر الكافي.

٣- الشكل (ج) يصور الحالة التي يكون عندها المقياس غير صادق وغير ثابت، لأن النقاط غير متمركزة حول المركز وغير متقاربة بل منتشرة في كل الاتجاهات.

٤- الشكل (د) يصور الحالة التي يكون عندها المقياس ثابتاً لكنه غير صادق. فقد يطبق المقياس أكثر من مرة ويعطي قراءات متقاربة (لاحظ تجمع النقاط) ولكنه لا يقيس ما صمم من أجله (لاحظ ابتعاد النقاط بشكل كبير عن المركز).



شكل رقم (١-٥)
رسم توضيحي لمفهوم العلاقة بين الصدق والثبات



*** تقارب النقاط او تجمعها في ماكن واحد يدل على ثبات الاستبانة**
***وقوع أو تمركز النقاط حول مركز الهدف يدل على صدق الاستبانة**

المصدر (بتصرف): http://explorable.com/research-methodology.

مكونات تقرير البحث العلمي:

بعد الانتهاء من الدراسة البحثية ينبغي توثيق جميع عناصرها فيما يسمى بتقرير البحث. ويبين الإطار أدناه المكونات التي يشتمل عليها تقرير البحث العلمي بوجه عام. وهذه المكونات قد تتفاوت في ترتيبها من دراسة لأخرى حسب طبيعة الدراسة والمنهجية المستخدمة فيها، (مراد وهادي ٢٠١٢، ٥٦٧). ويمكن للباحث الاستفادة منه في مرحلة إعداد تقرير البحث، كما يمكن استخدامه من قبل المحكمين كدليل إرشادي للتأكد من محتويات وترتيب مكونات البحوث.

- الجزء الأول - المحتويات:

١- صفحة عنوان الدراسة البحثية

٢- المحتويات

٣- قائمة الأشكال

٤- قائمة الجداول

- الجزء الثاني - التقرير:

- أولاً - المقدمة وتشمل:

١- تمهيد

٢- مشكلة الدراسة.

٣- هدف الدراسة.

٤- أهمية الدراسة.

٥- أسئلة و/ أو فرضيات الدراسة.

٦- مصطلحات الدراسة.

٧- حدود ومحددات الدراسة.

- ثانياً - أدبيات البحث ويشمل الإطار النظري والدراسات السابقة.

- ثالثاً - الإجراءات المنهجية:

١- وصف لمجتمع وعينة الدراسة.

٢- وصف لتصميم ومنهج الدراسة وأدوات جمع البيانات المستخدمة.

٣- صدق وثبات أدوات جمع البيانات.

٤- وصف وتبرير المعالجة والأساليب الإحصائية المستخدمة.

- رابعاً - النتائج:

عرض وتحليل البيانات لكل فروض وتساؤلات الدراسة.

- خامساً - الاستنتاجات والتوصيات.

- الجزء الثالث: المراجع.

- الجزء الرابع: الملاحق.

دور البرمجيات الحاسوبية في التحليل الإحصائي:

مع مرور الوقت ومن حين إلى آخر تستقبل الأسواق العديد والجديد من الحاسبات والبرامج الحاسوبية المختلفة والتي تتميز بقوتها وسرعتها وسهولة استخدامها. إن التحليل الإحصائي الغني بعدد كبير جداً من أدوات وأساليب التحليل قد حظي بأعداد مختلفة من البرامج الحاسوبية التي تساعد في تنفيذ تلك العمليات التحليلية بسرعة وسهولة وفاعلية كبيرة. ومن أمثلة برامج التحليل الإحصائي الشائعة الاستخدام برنامج SPSS, SAS, R, MINITAB, STATA وغيرها من البرامج العامة أو التخصصية بالإضافة إلى أنه يوجد تطبيق خاص بالتحليل الإحصائي ضمن برنامج EXCEL. إن إنتاج هذه البرامج ساعد بشكل كبير على تفادي وتجنب إجراء العمليات الحسابية والرسومات بواسطة اليد والتي بالطبع تحتاج إلى وقت طويل وجهود كبيرة فيما لو تم تنفيذها يدوياً، ومن ثم سهولة وقوع الخطأ في الحسابات وعدم الدقة في الرسومات البيانية. كما أن ذلك قد يؤدي بالباحث إلى استخدام أدوات تحليل غير ملائمة لكونها سهلة في الحساب (القحطاني، ١٤٣٢).

إضاءة إحصائية حول استخدام برامج التحليل الإحصائي:

ينبغي التنويه إلى أن برامج التحليل الإحصائي كغيرها من البرامج الأخرى ما هي إلا أداة أو وسيلة تساعد المستخدم في تنفيذ ما يحدده المستخدم، سواء كان الشيء المحدد صحيحاً أم خاطئاً، إن مخرجات التحليل تعتمد على ما تم اختياره وتحديده من قبل المستخدم وليس للبرنامج أي تدخل في اختيار أسلوب التحليل المناسب. وهذا يقتضي أهمية إلمام الباحث أو محلل البيانات بأداة (أسلوب) التحليل الإحصائي المراد تنفيذها من حيث استيفاؤها لشروط استخدامها، وهل هي مناسبة لتحليل المتغير أو المتغيرات محل الاهتمام؟ فمثلاً إذا أراد باحث في الشؤون الاجتماعية أن يتعرف على النزعة المركزية (المتوسط وليس المتوسط الحسابي) لمتغير الحالة الاجتماعية (أعزب، متزوج، أرمل، مطلق)، فإنه في هذه الحالة ينفذ من خلال برنامج التحليل الإحصائي المستخدم المنوال Mode وليس المتوسط الحسابي Mean أو الوسيط Median، وذلك كون المنوال هو الأداة الإحصائية الأنسب لحساب مركز بيانات متغير اسمي.

تعريف موجز ببرنامج التحليل الإحصائي SPSS:

بداية كانت الأحرف SPSS تمثل أوائل الكلمات في العبارة (Statistical Package for the Social Sciences) والتي تعني الحزم الإحصائية للعلوم الاجتماعية، ثم أصبحت فيما

بعد ترمز لأوائل الكلمات في العبارة (Statistical Product and Service Solutions) والتـي تعني الحلول الإحصائية للمنتجات والخدمات (Ho, 2006 ،11). وفي شـهر يناير من عام ٢٠١٠م قامت شـركة IBM الشـهيرة بشـراء برنامج SPSS من شركة SPSS، وبعـد امتلاك IBM للبرنامج تم تغيير المسـمى إلى IBM SPSS Statistics. وهنا سنكتفي باستخدام الرمز SPSS بدلاً من IBM SPSS Statistics.

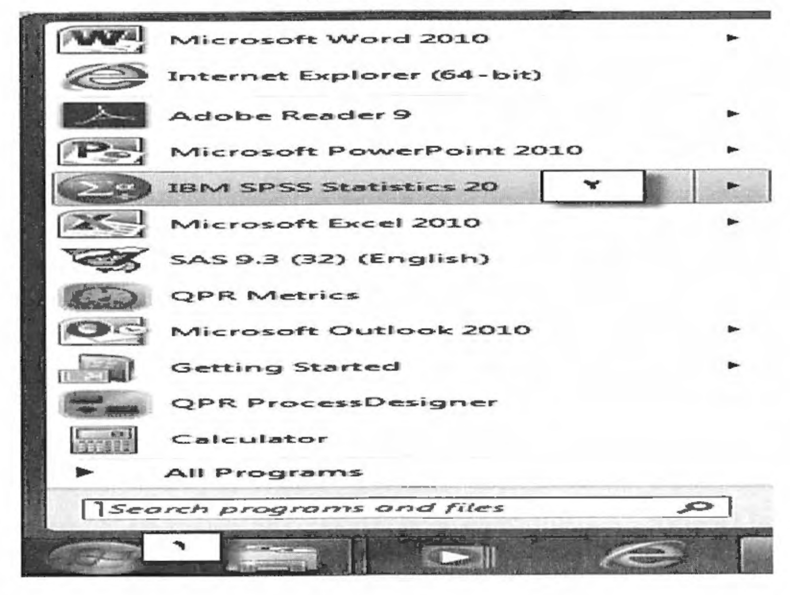
إن برنامج SPSS يعتبر من أكثر البرامج الإحصائية انتشاراً لسهولة استخدامه وكفاءته وفعاليته في التحليل الإحصائي وتحليل بيانات الدراسات الميدانية والبحوث العلمية.

وفيما يلي سـنتعرف وبشكل مختصر ومركز وبما يخدم أهداف الكتاب على أجزاء البرنامج وكيفية إدخال البيانات وإدارتها ومعالجتها ومن ثم تحليلها وصفياً واستدلالياً (القحطاني، ١٤٣٢).

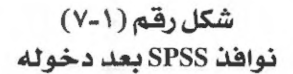
دخول (تشغيل) برنامج SPSS لبيئة النوافذ SPSS for Windows:

للدخول لبرنامج SPSS بوجه عام - وتحديداً IBM SPSS Statistics ٢٠ - تتبع الخطوات التالية حسب الترقيم التسلسلي من ١ إلى ٢ كما هو موضح بالشكل (١-٦) التالي:

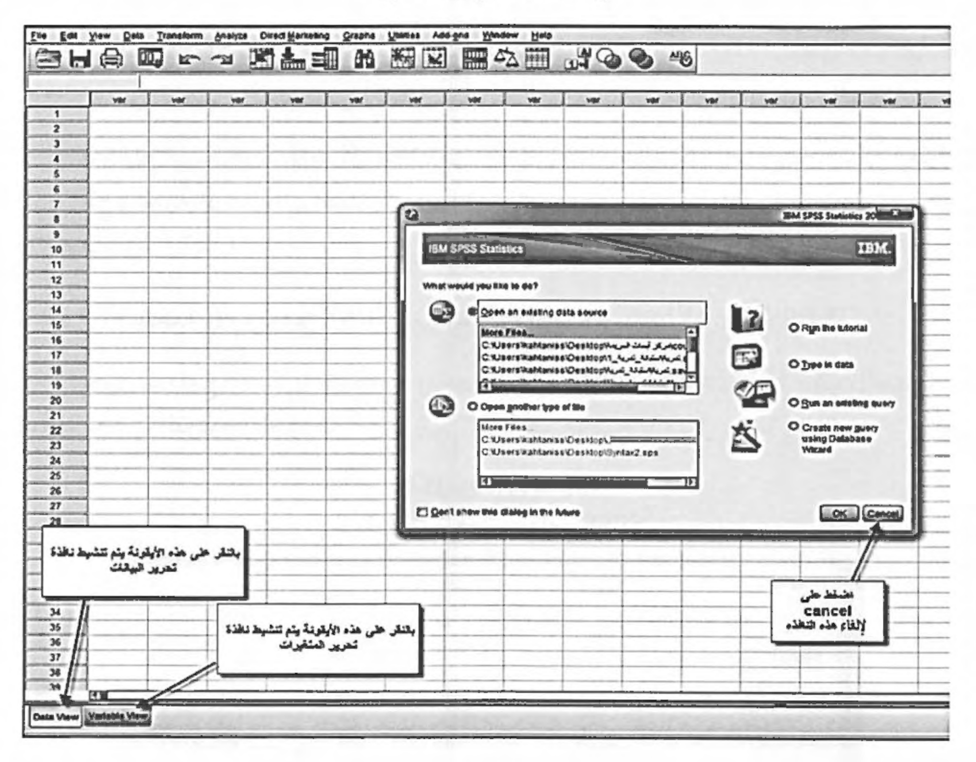
شكل رقم (١-٦)
خطوات تشغيل برنامج SPSS



وبالنقر على الأيقونة IBM SPSS Statistics 20 سوف يفتح لك البرنامج كما هو موضح في شكل (١-٧)



شكل رقم (١-٧)
نوافذ SPSS بعد دخوله



مكونات نافذتي تحرير المتغيرات والبيانات Variable View and Data View:

بعد فتح البرنامج تُفتح لك نافذتان: إحداهما تسمى نافذة تحرير المتغيرات Variable View لإدخال أسماء المتغيرات وتحديد خصائصها، والنافذة الأخرى تسمى نافذة تحرير البيانات Data View والتي يتم من خلالها إدخال البيانات وإدارتها وإجراء التعديلات المطلوبة على المتغيرات وتنفيذ التحليلات الإحصائية.

قبل البدء في إدخال البيانات نقوم أولاً بتعريف متغيرات الدراسة، وذلك بتحديد أسمائها ونوعها ودرجة قياسها، وغير ذلك من خلال نافذة تحرير المتغيرات Variable View .

والشــكل (١-٨) التالي يمثل نافذة تحرير المتغيرات وما تحتويه من أيقونات لتســمية وتحديد خصائص المتغيرات.

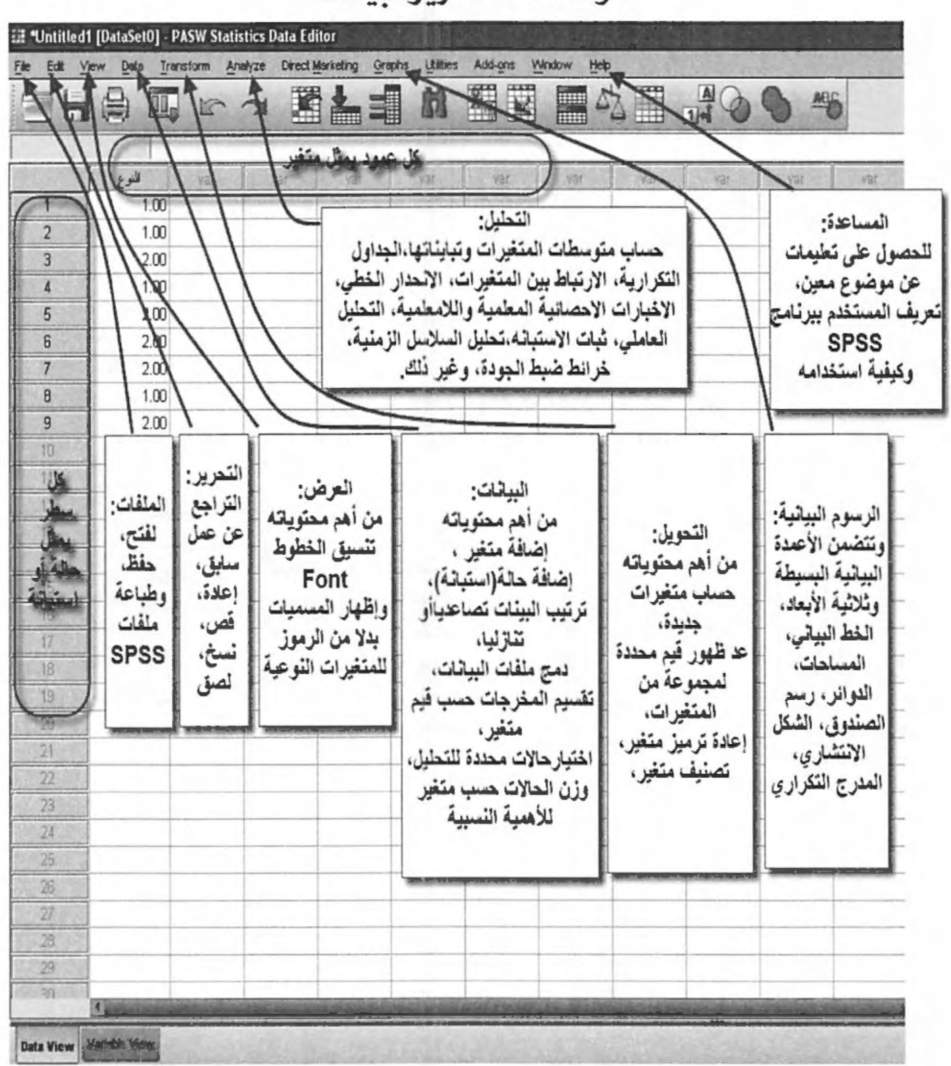
شكل رقم (١-٨)
مكونات نافذة تحرير المتغيرات



وبعد تعريف المتغيرات، نبدأ بإدخال بيانات كل متغير من تلك المتغيرات التي تم تعريفها في مرحلة تعريف متغيرات الدراسة. والشكل (١-٩) التالي يوضح بشكل مختصر مكونات نافذة تحرير البيانات وما تحويه من قوائم لإدارة البيانات، إجراء التعديلات اللازمة على المتغيرات، إدراج الرسومات البيانية، التحليل الإحصائي، ... إلخ.

شكل رقم (١-٩)

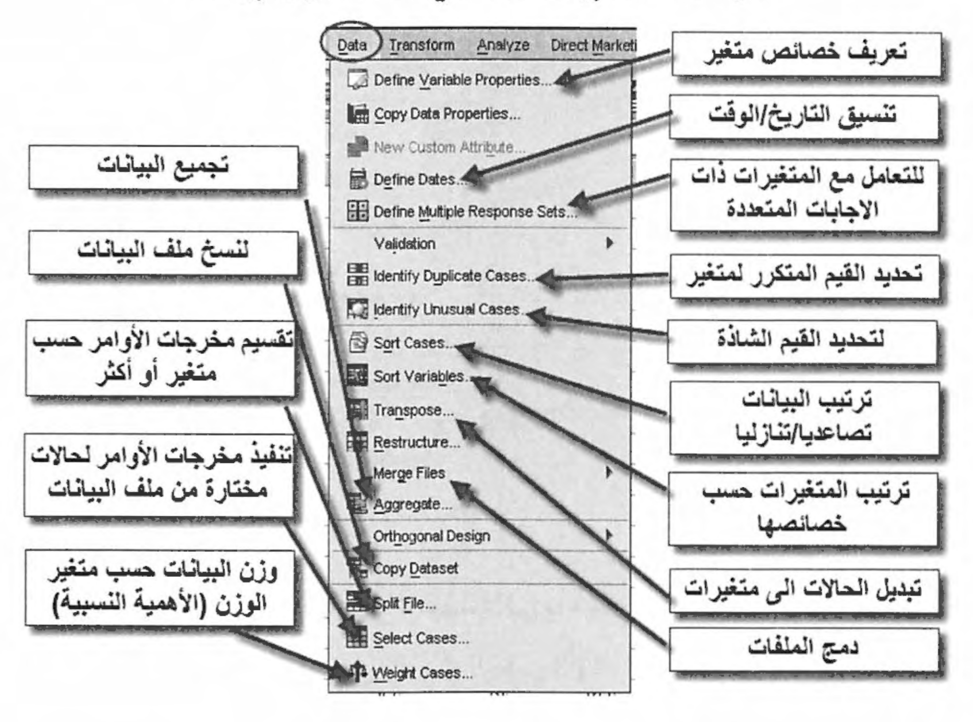
مكونات نافذة تحرير البيانات



إدارة البيانات والمتغيرات ومعالجتها باستخدام SPSS:

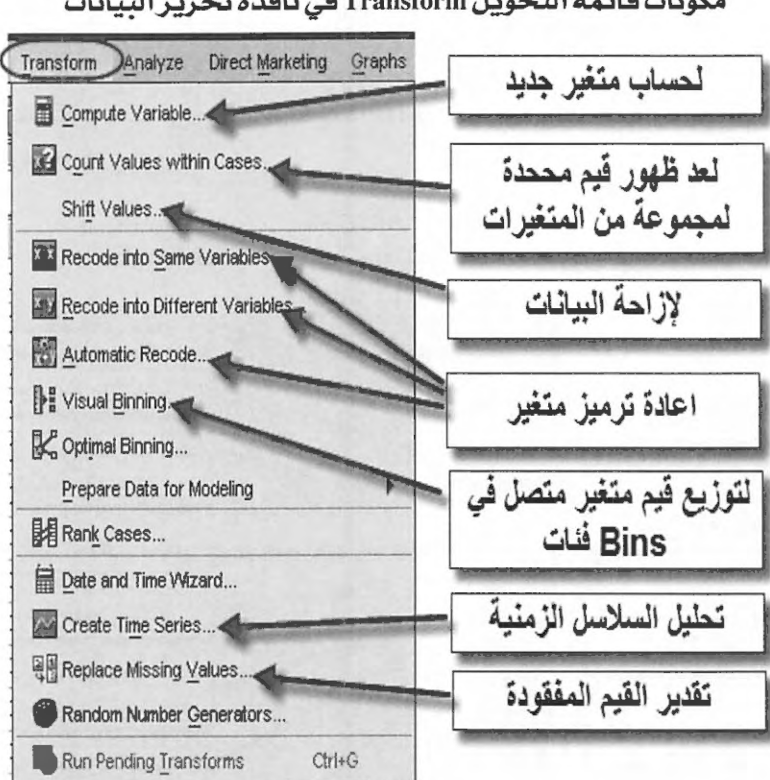
إن قائمتي البيانات Data والتحويل Transform تتضمنان العديد من الأوامر التي تستخدم لإدارة البيانات وإجراء التعديلات اللازمة على المتغيرات. وفيما يلي سوف نعرض بشكل مختصر أهم محتويات هاتين القائمتين كما هو موضح في الشكلين (١-١٠) و(١-١١).

شكل رقم (١-١٠)
مكونات قائمة البيانات Data في نافذة تحرير البيانات



شكل رقم (١-١١)

مكونات قائمة التحويل Transform في نافذة تحرير البيانات

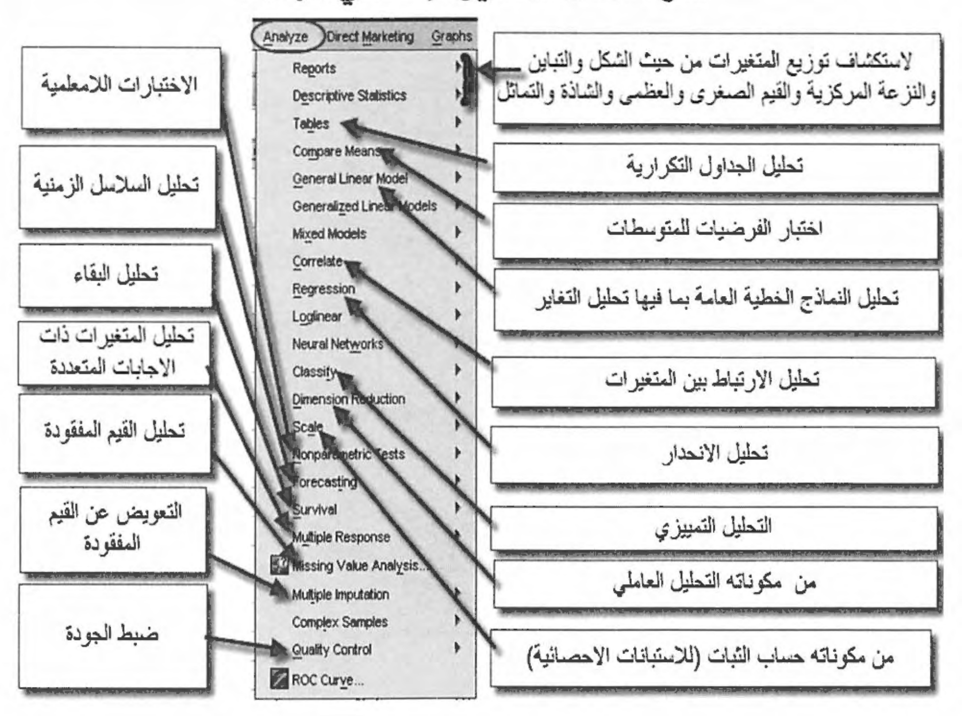


التحليل الإحصائي الوصفي والاستدلالي باستخدام SPSS:

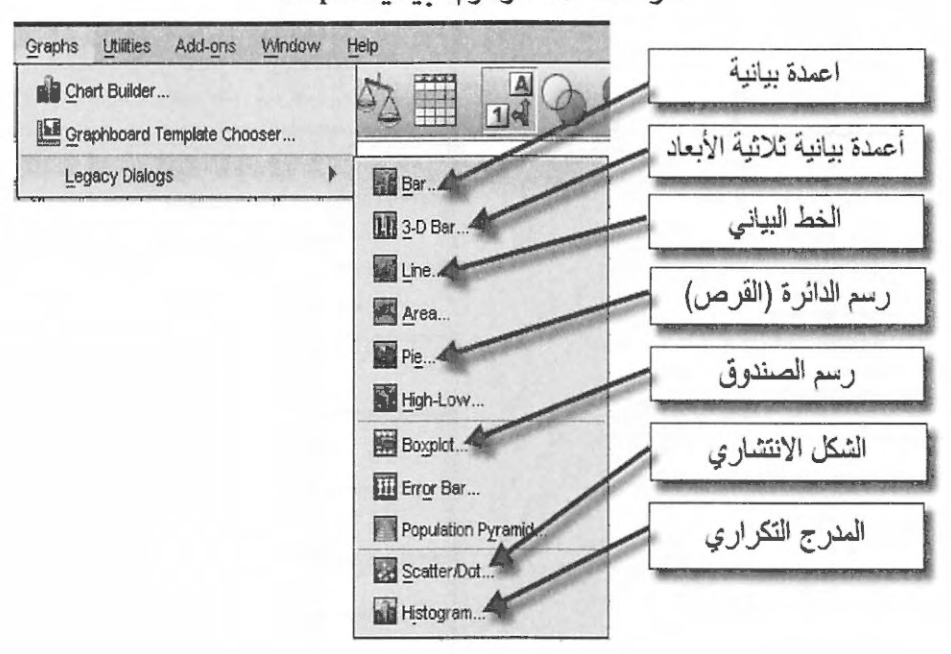
إن قائمتي التحليل Analyze والرسم البياني Graph تستخدمان لأغراض التحليل الإحصائي المطلوب بعد إدخال البيانات وتعريف المتغيرات ومعالجتها، سواء بحساب متغيرات جديدة أو تحويل متغيرات من صيغ معينة إلى أخرى أو غير ذلك حسب احتياج الباحث. وفيما يلي توضيح لبعض محتويات هاتين القائمتين كما هو موضح بالشكلين (١-١٢) و(١-١٣) التاليين:

شكل رقم (١-١٢)

مكونات قائمة التحليل الإحصائي Analyze



شكل رقم (١-١٣)
مكونات قائمة الرسوم البيانية Graphs



الفصل الثاني
الإحصاء: تعريفه، أهميته، ومفاهيمه الأساسية
Statistics: Its Definition, Importnace, and Basic Concepts

تعريف الإحصاء:

يعرف الإحصاء على أنه مجموعة الأساليب والطرق والإجراءات الرياضية التي تهتم بجمع وتحليل البيانات بهدف الوصف والاستنتاج والتنبؤ والتحكم. فالإحصاء علم وفن ووسيلة للاستدلال على العام من الخاص وإعطاء قرارات حكيمة في مواجهة حالات عدم التأكد. وكما يلاحظ من التعريف يوفر الإحصاء لنا الأساليب والأدوات التالية:

١- أساليب وأدوات التصميم والتخطيط لآلية جمع البيانات للدراسات والبحوث.

٢- أساليب وأدوات تنظيم وتلخيص ووصف البيانات.

٣- أساليب وأدوات الاستدلال والتحكم والتنبؤ الإحصائي من البيانات (Agresti & Finlay, 2009).

أهمية الإحصاء:

إن علم الإحصاء يعد ركيزة هامة من ركائز التخطيط التنموي السليم والتطوير المستمر، فهو العلم الذي يدخل ويستخدم تقريباً في جميع المجالات والتخصصات المختلفة الاجتماعية منها والتربوية والاقتصادية والإدارية والصحية والسياسية والبيئية والزراعية والصناعية ... إلخ. إنه الأسلوب العلمي في التحري عن حقيقة الظواهر المختلفة وتقصيها ودراستها بعناية، وذلك في سبيل وصفها بالشكل الصحيح واستخلاص النتائج عنها، وهو يعتبر من أدق وسائل البحث العلمي المرتكز على الأسلوب والنظرية. إن للإحصاء مهام عدة منها جمع البيانات، تحليل البيانات كمياً وبيانياً، صياغة الفرضيات حول ظاهرة معينة ووضع الاختبارات المناسبة لها، استخلاص النتائج، اتخاذ القرارات المناسبة اتجاهها والتنبؤ بالقراءات المستقبلية حولها. كما أن الإحصاء يعد حلقة هامة في عمليات دراسة وضبط الجودة للمنتجات الخدمية والصناعية على حد سواء. لقد بدأ استخدام الإحصاء منذ العصور القديمة وأخذ في التطور في أساليبه وطرقه ونظرياته حتى أصبح من أهم العلوم التي ينبغي أن يعتمد عليها في بناء المجتمعات وتطويرها.

ونظـراً لمـا احتله علم الإحصـاء من مكانة هامـة بـين الأمم وإدراك الحاجة إليه، خصصت له أقسـام في الجامعات والكليات والمعاهد لدراسته باعتباره علماً والتخصص فيه، واسـتحدثت أقسام ومراكز لجمع البيانات وتحليلها في مختلف المؤسسات الخاصة والعامة وأنشـأت من أجله الجمعيـات والمجلات العلمية وأقيمت لـه المؤتمرات العلمية والـدورات التدريبية وورش العمـل. وإدراكاً للدور الفعال للإحصـاء والتوعية بأهميته وإسـهاماته المختلفة والعديدة في مختلف مجالات الحياة فقد خصص يوم ٢٠ من شهر أكتوبر عام ٢٠١٠ كأول يوم عالمي للإحصاء والذي أقرته اللجنة الإحصائية للأمم المتحدة (UNSD) وتم الاتفاق على أن يحتفل بالإحصاء في مثل هذا اليوم كل خمسة أعوام.

إن مدى نجاح الخطط التنموية والتطوير والتحسـين المسـتمر - ولاسيما في ظل عصر التغيير السـريع والانفجار السـكاني وتعدد واختلاف الموارد البشرية والمادية - يقاس بما تسـتند إليه من بيانات ومعلومات ومؤشرات إحصائية صحيحة ودقيقة عن مختلف المتغيرات الاجتماعية والديموغرافية والاقتصادية والسياسية ... إلخ. ولا بد أن تكون هـذه البيانات والقراءات الإحصائية متجددة وموائمة لكل ما يسـتجد على المستوى المحلي والإقليمي والعالمي بما يلبي احتياجات العملية التنموية.

إن أسـلوب الاعتمـاد على الآراء الشـخصية والاجتهاد والارتجاليـة في عمليات التخطيط والتطويـر واتخاذ القرارات بعيداً عن لغة الأرقام والقياس الكمي والتحليل سـيؤدي حتماً إلى الفشـل للوصول إلى المسـتوى المأمول من التنمية وبطء في عملية التطوير والتحسـين المسـتمر. ولكي يتم أيضـاً التخطيط والتطوير بشـكل صحيح والوصول إلى قرارات سـليمة لابد أن تكون البيانـات والمعلومات الإحصائية التي تم الاسـتناد إليها صحيحة وذات دقة وذات جودة عالية لأن ما بني على باطل فهو باطل، وهـذا بدوره يقودنا إلى قضية الاهتمام بعملية الحصـول على البيانات وجمعها وفق طرق إحصائية علمية صحيحة.

إن الشـفافية والمصداقيـة والنزاهة والوعـي بأهمية إتاحة البيانـات والمعلومات والمؤشرات الإحصائية الصحيحة سواء على مستوى الأفراد أو على مستوى المؤسسات والمنظمات بشـقيها الخاص والعام يوفر وبلا شـك مصدراً جيداً للبيانات والمعلومات التي تلبـي احتياجات المخططين والمطورين وأصحاب القـرار والباحثين للوصول إلى الخطط التنموية السليمة واتخاذ القرارات الصحيحة وإيجاد الحلول للقضايا المختلفة والتي من شأنها الإسهام في بناء المجتمع والرقي بحياة الأفراد على مختلف الأصعدة التعليمية والاقتصادية والصحية والأمنية.

أنواع المتغيرات Variables:

يعرف المتغير على أنه خاصية أو سمة يمكنها أن تأخذ قيماً مختلفة أو متغيرة بين وحدات الدراسة في العينة أو المجتمع (Agresti & Finlay, 2009, 11). وتنقسم المتغيرات بوجه عام حسب نوع بياناتها إلى قسمين رئيسيين:

١- متغيرات كمية Quantitative Variables: وهي المتغيرات التي يتم قياسها كمياً، وقيمها عبارة عن أرقام كمية، مثل العمر، الوزن، الطول، الدخل، ... إلخ.

وهذا النوع من المتغيرات بدوره يمكن تقسيمه إلى قسمين فرعيين:

١- متغيرات متصلة Continuous Variables: وفي هذا النوع من المتغيرات يوجد عدد لانهائي من القيم بين أي قيمتين متجاورتين، مثل الطول، الوزن، ... إلخ.

٢- متغيرات منفصلة أو متقطعة Discrete Variables: وفي هذا النوع من المتغيرات تكون البيانات قابلة للعد ومنفصلة بعضها عن بعض بحيث يوجد فجوات Gaps بين القيم، مثل عدد أفراد الأسرة، عدد المواليد، ... إلخ. وتجدر الإشارة إلى أن البيانات المتقطعة أو المنفصلة قد تكون بيانات أو أرقاماً كسرية ولكن يوجد بينها فجوات، فهناك حالات وتطبيقات عملية كثيرة لها مثل درجات الطلاب فقد تأخذ كسوراً منتهية مثل ربع، نصف، ثلاثة أرباع، ومن ثم تكون الدرجات مثلاً كالتالي: ١٢,٢٥، ١٠,٥٠، ١٤,٧٥، ١٣، ...

٢- متغيرات نوعية Categorical Variables: وهي التي لا يمكن التعبير عنها كمياً بأرقام، بل يمكن وصفها حسب صفات أو أسماء أو رموز مختلفة، مثل الجنسية (سعودي، غير سعودي)، المؤهل العلمي (ابتدائي، متوسط، ثانوي، ...)، ... إلخ، (فهمي، ٢٠٠٥، ٢٢-٢٣).

المتغيرات المستقلة والتابعة Independent and Dependent Variables:

في كثير من أدوات التحليل الإحصائي كتحليل الانحدار الخطي وتحليل التباين مثلاً ينبغي التمييز بين نوعين من المتغيرات حسب دورها في التحليل الإحصائي، وتعرف بالمتغيرات المستقلة والمتغيرات التابعة.

حيث يعرف **المتغير التابع** (الاستجابة أو الأثر) على أنه المتغير الذي يسعى الباحث إلى دراسته وتفسير تباين قيمه و/أوالتنبؤ بقيمة من قيم المتغيرات المستقلة.

في حين يعرف المتغير المستقل (السبب أو المؤثر) على أنه ذلك المتغير الذي يؤثر في المتغير التابع ويفسر التباين فيه.

البناء Construct: هو مفهوم تجريدي لوصف أو تفسير ظاهرة معينة، مثل الولاء للمنظمة، القيادة التحويلية، الذكاء، التوتر، الطول، الوزن، الدخل، ... إلخ.

التعريف الإجرائي للبناء Operational Definition: هو الوصف الدقيق والواضح لمعنى البناء وكيفية قياسه عملياً بحيث يمكن لشخص آخر فهم معنى البناء وإمكانية قياسه مرة أخرى بدقة. ففي البحوث العلمية ينبغي تضمين التعريف الإجرائي للمفاهيم البنائية المستخدمة في البحوث بحيث يسهل توصيل المعنى من باحث لآخر. فالدخل مثلاً قد يختلف مفهومه وقياسه من شخص إلى آخر حسب الدراسة أو البحث، فقد يقصد به الدخل السنوي للأسرة بالريال السعودي في دراسة معينة، في حين يعرف على أنه الدخل الشهري للفرد بالدولار في دراسة أخرى، وقس على ذلك غيره من الأمثلة: مثل الإبداع، الموهبة، الحالة الاقتصادية، ... إلخ.

مستويات قياس المتغيرات Variables Measurement Levels:

تكمن أهمية مستوى قياس المتغير (كذلك الحال بالنسبة لأنواع المتغيرات) في تحديد اختيار الأداة الإحصائية المناسبة في تحليل البيانات. وقد قسم ستيفنز Stevens (1946) مستويات قياس المتغيرات إلى أربعة أقسام مرتبة حسب درجة قوة المقياس من الأضعف إلى الأقوى كالتالي:

١- متغيرات اسمية Nominal Variables:

- قيم هذه المتغيرات عبارة عن ألفاظ أو رموز.
- قد تكون قيم المتغيرات الاسمية أرقاماً لكنها لا تتضمن المعنى الكمي لها مثل أرقام الهواتف.
- لا يمكن ترتيب قيمها تصاعدياً أو تنازلياً.
- عدم ملاءمة إجراء العمليات الحسابية مثل الجمع الطرح وغيرها عليها.
- أمثلة: الجنسية، النوع (ذكر / أنثى)، أرقام لوحات السيارات، أرقام الهوية الوطنية.

٢- متغيرات رتبية Ordinal Variables:

- قيم المتغير الرتبي عبارة أرقام أو ألفاظ أو رموز.

- هذا المقياس أعلى في مستوى القياس من الاسمي.
- يمكن ترتيب قيمه تصاعدياً أو تنازلياً، أو من الأفضل إلى الأسوأ أو العكس، ... إلخ.
- عادة تعطى قيم هذه المتغيرات أرقاماً متوالية تعبر عن الترتيب والأفضلية.
- أمثلة: المؤهل العلمي، المرتبة الوظيفية، مقياس ليكرت للاتجاهات.

٣- متغيرات فترية Interval Variables:

- وهو أقوى من الاسمية والرتبية.
- قيمه عبارة عن أرقام فقط تحمل المعنى الكمي لها.
- يمكن ترتيب قيمه تصاعدياً أو تنازلياً.
- الفرق بين القيم المتتالية لهذا النوع من المتغيرات متساوية، وهذا غير متحقق في المتغيرات الرتبية، ولذا سمي بالمقياس الفتري. فمثلاً نجد التغير في ارتفاع درجة الحرارة في إحدى المدن من ٣٠ إلى ٣٥ درجة مئوية يساوي نفس التغير الحاصل في مدينة أخرى ارتفعت درجت حرارتها من ٣٥ إلى ٤٠ درجة مئوية.
- الصفر في هذا النوع من المتغيرات يعتبر نسبياً أي لا يعني الصفر المطلق بمفهومه العام «لا شيء»، وإنما الصفر هنا يعبر عن حالة محددة للمتغير. فدرجة حرارة الجو تساوي صفراً درجة مؤية لا يعني اختفاء درجة الحرارة وإنما تعبر عن حالة محددة للطقس.
- كما أنه يمكن تطبيق العمليات الحسابية عليه عدا الضرب والقسمة لأن الصفر هنا غير مطلق وغير حقيقي. فمثلاً درجة الحرارة ٦٠ لا تساوي ضعف درجة الحرارة ٣٠ درجة مئوية.
- أمثلة: درجة الحرارة، معدلات الذكاء، والوقت.

٤- المتغيرات النسبية Ratio Variables:

- أعلى من مستويات القياس الثلاثة السابقة.
- يحمل جميع خصائص المتغيرات الفترية عدا خاصية الصفر المطلق والتي تعني هنا «لا شيء» أو اختفاء أو انعدام الشيء.
- كذلك يمكن تطبيق جميع العمليات الحسابية عليه.
- أمثلة: الدخل، العمر، الوزن، عدد أفراد الأسرة.

إضاءة إحصائية حول مستويات قياس المتغيرات في التحليل الإحصائي:

فـــي كثير من الأبحاث يتم معاملة المتغيرات الرتبيـــة أثناء التحليل الإحصائي على اعتبار أنها متغيرات فترية، وذلك لسهولة ومألوفية الأساليب الإحصائية المعتمدة على المتغيرات الفترية مثل المتوسط الحسابي والانحراف المعياري واختبارات t، وغيرها من أســـاليب التحليل الإحصائي المعلمي. ومن الأمثلة الشائعة على ذلك معاملة المتغيرات المقاسة بمقياس ليكرت للاتجاهات على أنها متغيرات فترية أثناء التحليل الإحصائي. لقد استعرض وارنر Warner (2008) النقاشات التي تناولها بعض الباحثين حول هذا الموضوع بين ممانع وغير ممانع لاســـتخدام الأســـاليب المعلمية في حال كون البيانات رتبية. وقد خلص وارنر إلى أنه يمكن اســـتخدام الأســـاليب المعلمية في حال البيانات الرتبيـــة على الرغم من أن بعض الباحثين والأكاديميين والمجلات العلمية لديها تحفظ حول قبـــول ذلك. وكان لـ Tabachnick & Fidell (2007) وغيرهم من الباحثين رأي حول هذا الموضوع حيث يرون أن توزيع المتغير الكمي أهم من مســـتوى قياســـه رتبي أو فتري أو نســـبي في اختيار الأســـاليب المعلمية أو اللامعلمية فـــي تحليل البيانات. فكـــون توزيع المتغير طبيعياً أو قريباً مـــن الطبيعي، أو متماثلاً أو غير متماثل أهم من كون مســـتوى قياس المتغيـــر رتبياً أو فترياً في الاختيار بين أســـاليب التحليل المعلمية واللامعلمية. وهذا مقبول عند كثير من الباحثين عند توفر الحد الأدنى لحجم العينة المطلوب مع التأكد من أن توزيع المتغيرات الرتبية متماثل. ويفضل المؤلف عدم معاملة البيانات الرتبية في التحليل الإحصائي على اعتبار أنها فترية أو نسبية لعدة أسباب. أولاً، عدم وجود ما يمنع أو يعوق اســـتخدام أساليب التحليل الإحصائية (اللامعلمية) المخصصة للبيانات الرتبية، ولاسيما في ظل توفر برامج التحليل الإحصائي المتعددة. واســـتخدام كثير من الباحثين للأســـاليب المعلمية في حال البيانـــات الرتبية لا يعني بالضرورة أنها أفضل أو أصح بل جرت العادة على اســـتخدام تلك الأســـاليب وتناقلها الباحثـــون عبر الزمن، وذلك لأن الأســـاليب المعلمية تم تطويرهـــا أولاً ثم أتى بعدها تطوير الأســـاليب اللامعلمية والتي لم تكن متاحة في زمن الأســـاليب المعلمية، وذلك لمعالجة الخلل في الأســـاليب المعلمية عند عدم تحقق بعض شروط صلاحيتها. ثانياً، عـــدم تحقق التماثل فـــي توزيع البيانات فـــي كثير من الأحيان ولاســـيما مع العينات صغيـــرة الحجم. ثالثاً، قيم المتغيـــرات الرتبية تعكس فقط ترتيب القيم، والفروق بين تلـــك القيم لا تعكس حجم الفرق. فمثـــلاً إذا كان لدينا متغير ما مقاس على مقياس ليكرت الخماسي ذو القيم «موافق بشدة (٥)، موافق (٤)، محايد (٣)، غير موافق (٢)،

غير موافق بشــدة (١)» فإن الفرق بين موافق بشــدة (٥) وموافق (٤) لا يساوي الفرق بين غير موافق (٢) وغير موافق بشــدة (١) وليس له معنى، وبإمكان القارئ إســقاط نفس المفهوم على اختبارات الفروق بين المتوســطات لمجموعتين أو أكثر. وقد أشــار إلى هذا الجانب كثير من المهتمين حيث أشــار Huck (2012) إلى أن أحد الأســباب لاستخدام أساليب التحليل الإحصائي اللامعلمي هو كون قياس المتغير رتبياً.

في برنامــج التحليل الإحصائي SPSS، تصنف مســتويات قيــاس المتغيرات إلى ثلاثة أقســام: Nominal للمتغيرات الاســمية، Ordinal للمتغيرات الرتبية، وScale للمتغيرات الفترية والنســبية. فيلاحظ أنه تم دمج المقياســين الفتري والنسبي معاً، وذلك لعدم الحاجة لهذا التصنيف في التحليلات الإحصائية المألوفة.

المجتمع الإحصائي Population Statistical:

يعــرف على أنه تجمع من الأشــياء أو الأفراد والتي تشــترك في خصائص معينة ويجمعها إطار واحد. فمثلاً عند دراســة مشكلة «قياس رضا المستفيدين من خدمات الأحــوال المدنية فــي مدينة الرياض» فإن مجتمع الدراســة هنا هو المســتفيدون من خدمات الأحوال المدنية في مدينة الرياض.

المجتمع المستهدف Target Population:

هو المجتمع الذي يرغب الباحث دراسته وتعميم النتائج عليه.

مجتمع الدراسة Study Population:

هــو المجتمع الممكــن الدخول إليه والوصــول إلى أي مفردة من مفرداته وســحب العينــة منه، ومن ثم يتم تعميم نتائج تحليل العينة عليه، فمثلاً إذا أراد باحث دراســة «أثر تطبيق الحكومة الإلكترونية على الحد من ممارســات الفساد الإداري والمالي في المؤسســات الحكومية في المملكة العربية الســعودية»، فإن المجتمع المستهدف في هذه الحالــة هو جميع المؤسســات الحكومية في المملكة العربية الســعودية. فإذا اقتصر الباحث في ســحب عينة دراسته من المؤسســات الحكومية في مدينة الرياض فقط، فإنه في هذه الحالة يصبح مجتمعاً دراســته المؤسســات الحكومية في مدينة الرياض فقط ويسمى مجتمع الدراسة، ومن ثم سيتم تعميم نتائج دراسته على مجتمع الدراسة (المؤسســات الحكومية فــي مدينة الرياض) ولا يصح له أن يعمم نتائج دراســته على

المجتمع المستهدف (المؤسسات الحكومية في السعودية). وتجدر الإشارة الى أنه يمكن تعميم نتائج الدراسة على المجتمع المستهدف إذا كان مجتمع الدراسة ممثلاً تمثيلاً جيداً له.

معلمة (مؤشر) المجتمع Population Parameter:

إحدى خصائص مجتمع الدراسة التي يتم قياسها كمياً في حال حصر جميع مفردات مجتمع الدراسة كالمتوسط الحسابي والنسبة. فمثلاً يمكننا معرفة متوسط أعمار المتدربين في معهد الإدارة العامة بالرياض لعام ١٤٣٣هـ، وذلك من خلال قاعدة البيانات المتوفرة لدى إدارة القبول والتسجيل بالمعهد.

تحليل البيانات Data Analysis:

إن تحليل البيانات يعد جزءاً هاماً ومكوناً أساسياً في الدراسات والبحوث حيث يساعد الباحث في الإجابة عن أسئلة وفرضيات دراسته أو بحثه وتحقيق أهدافها وفهم الظاهرة محل الدراسة وتفسيرها والوصول إلى استنتاجات ذات قيمة علمية وعملية. فبعد مرحلة جمع البيانات يحصل الباحث على كم كبير من البيانات، وهذه البيانات لن تكون ذات فائدة أو عديمة الجدوى إذا لم يتم معالجتها وتلخيصها وتحليلها وتفسيرها في ضوء أهداف وتساؤلات البحث وفرضياته. وينقسم تحليل البيانات حسب طبيعة البحوث والبيانات إلى نوعين رئيسيين هما التحليل الكمي والتحليل الكيفي.

أولاً - التحليل الكيفي Qualitative Analysis:

إن البيانات الناتجة عن الدراسات والبحوث الكيفية تكون غالباً على شكل ملاحظات وتعليقات وآراء مكتوبة أو مشاهدة أو مسموعة وخصائص وصفات ورموزاً مفصلة يتم الحصول عليها من وحدات الدراسة. وتحليل هذا النوع من البيانات غير الرقمية (غير الكمية) لتحقيق أهداف الدراسة أو البحث وتفسير الظاهرة محل الدراسة وفهمها يسمى بالتحليل الكيفي. ويختلف تحليل البيانات الكيفية عن البيانات الكمية في عدة جوانب. والجدول (٢-١) يوضح أبرز الاختلافات بين التحليل الكمي والكيفي (Bhattacherjee, 2012; Johnson, Christensen, 2012).

جدول رقم (٢-١)
أبرز الاختلافات بين التحليل الكمي والكيفي

التحليل الكمي	التحليل الكيفي
يعرف الباحث مقدماً وبوضوح ما يبحث عنه بشكل محدد من خلال تحديد أسئلة الدراسة وفرضياتها.	مع أن هناك هدفاً عاماً للدراسة إلا أن الباحث لا يعرف مقدماً بوضوح تام ما يبحث عنه بشكل محدد، فالعملية هنا استكشافية بحته، فخلال دراسته تطرأ تساؤلات جديدة يسعى للإجابة عنها.
يتم التركيز على أدبيات الدراسة أو البحث قبل البدء في تحليل البيانات.	يتم التركيز على أدبيات الدراسة أو البحث في مرحلة التحليل.
يتم تحليل البيانات وحساب المؤشرات الإحصائية الكمية والرسومات البيانية وبناء النماذج الإحصائية واختبارات الفرضيات لتفسير ما تم مشاهدته حول الظاهرة محل الدراسة.	يهدف التحليل إلى تقديم وصف كامل ومفصل حول الظاهرة محل الدراسة.
يستخدم الباحث أدوات جمع بيانات مثل الاستبانة.	هنا يعتبر الباحث هو أداة جمع البيانات حيث يلاحظ بنفسه ويسجل البيانات ويقرأ الوثائق والصور ويستمع ويسجل تعليقات وأقوال أفراد الدراسة.
تكون البيانات غالباً ذات طابع كمي مثل الأرقام والإحصائيات.	تأخذ البيانات الطابع الكيفي (غير الكمي) مثل الكلمات والصور والملفات الصوتية وغير ذلك.
يعتمد بشكل كبير على الأساليب الإحصائية مع الاستقلالية الكبيرة عن تأثير الباحث.	يعتمد على القدرة التحليلية للباحث وعلى مدى فهمه واستيعابه للظاهرة محل الدراسة وأبعادها المتعددة ومعرفته بالمجتمع الذي جمعت منه البيانات.
هناك فاصل زمني بين جمع البيانات وتحليلها، حيث يتم جمع البيانات أولاً ثم يأتي بعد ذلك مرحلة تحليل البيانات.	ليس هناك فاصل زمني محدد بين جمع البيانات وتحليلها فعملية التحليل لا تترك حتى الانتهاء من جمع البيانات وإنما تبدأ منذ لحظة البدء وأثناء فترة جمع البيانات وتستمر حتى بعد جمع البيانات، فعملية جمع البيانات والتحليل مستمرة ومتداخلة ويكمل بعضها بعضاً.

وفـــي الملاحق، يتضمن الملحق (٣) نبذة مختصـــرة عن تحليل البيانات في البحوث الكيفية تشمل مراحل تحليل البيانات الكيفية وأهم البرامج الحاسوبية المستخدمة في تنظيمها وتحليلها.

ثانياً - التحليل الكمي Quantitative Analysis:

إن مفهوم التحليل الكمي للبيانات يشـــير إلى ذلك التحليل الذي يتخذ من الأرقام والعمليات والأســـاليب الرياضية أساســـاً له. والتحليل الكمي يهدف إلى تحليل تلك البيانـــات التي يتم الحصول عليها من الدراســـات والبحوث الكمية. وفي هذا الكتاب ســـيتم تناول أســـاليب التحليل الإحصائي باعتباره أحد أهم أســـاليب التحليل الكمي للبيانات. إن أســـاليب التحليل الإحصائي بشـــكل عام تنقســـم إلى قسمين رئيسيين (انظر الشكل ٢-١) هما:

١- أساليب التحليل الإحصائي الوصفية Descriptive Statistical Analysis:

عبـــارة عن مجموعة مـــن الأدوات الرياضية تهدف إلى تنظيـــم وتلخيص البيانات للوصـــول إلى معلومات ذات معنى وقيمة. فهو يســـاعد علـــى التعرف على خصائص الظاهرة محل الدراســـة وســـلوكيات توزيع بياناتها وعلاقة متغيراتها وأبعادها بعضها ببعض وعلاقتها أيضاً بظواهر أخرى. وهو يســـتخدم دائماً في الدراســـات والبحوث والممارسات الحياتية اليومية. وسيتم التعرف على أساليب التحليل الإحصائي الوصفي بشيء من التفصيل في الفصل الرابع.

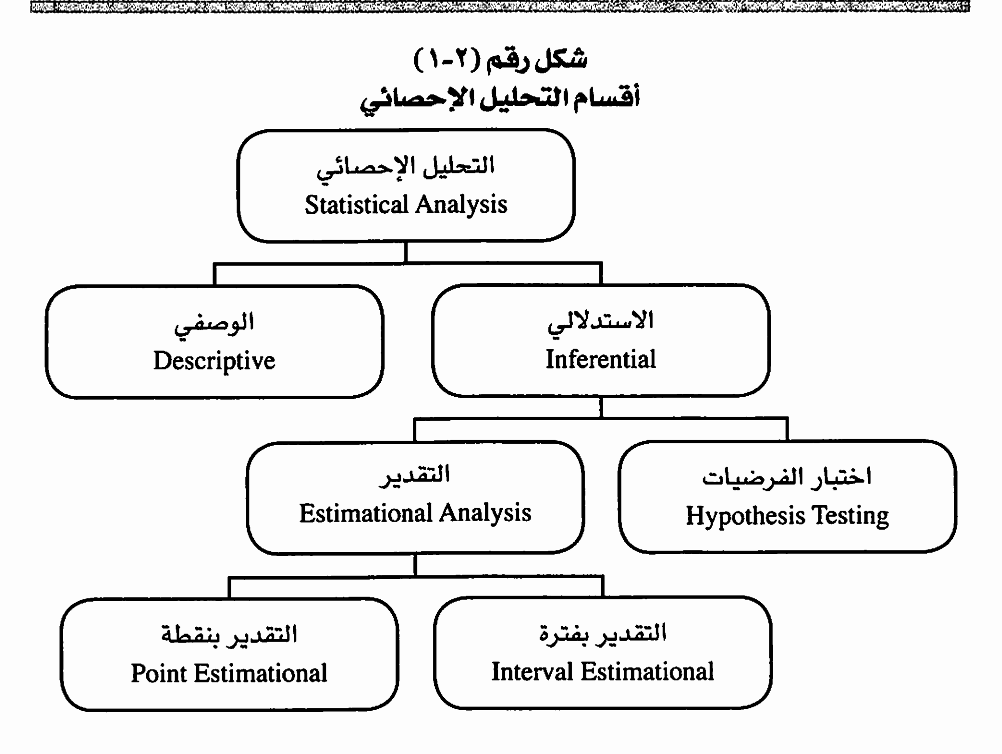
٢- أساليب التحليل الإحصائي الاستدلالية Inferential Statistical Analysis:

عبـــارة عن مجموعة من الأدوات الرياضية تهدف إلى الوصول إلى نتائج معينة من خلال تحليل بيانات العينة يمكن تعميمها على مجتمع الدراســـة. ومن اســـتخدامات الأســـاليب الإحصائية الاســـتدلالية في الدراسات البحثية دراســـة جوهرية (معنوية) العلاقـــة بين المتغيرات، جوهريـــة الفروق بين مجموعتين أو أكثـــر، أو التقدير والتنبؤ (بري، هندي وراضي، ١٩٩٨، ٥).

وتنقسم أساليب التحليل الإحصائي الاستدلالية إلى قسمين رئيسيين هما:

- التقدير بنقطة والتقدير بفترة Point and Interval Estimation.
- الاختبارات المعنوية (الاختبارات الاحصائية) Significance (Statistical) Tests.

شكل رقم (٢-١)
أقسام التحليل الإحصائي



ويمكن تصنيف أساليب التحليل الإحصائي الاستدلالية حسب عدد المتغيرات التابعة إلى قسمين هما:

١- أساليب التحليل الإحصائي للمتغيرات الأحادية Univariate Statistical Analysis، وفي هذه الحالة يكون لدينا متغير تابع واحد فقط وصفر أو أكثر من المتغيرات المستقلة. وسيتم لاحقاً مناقشة بعض من هذه الأساليب والتي تعتبر أكثر استخداماً في الدراسات والبحوث.

٢- أساليب التحليل الإحصائي للمتغيرات المتعددة Multivariate Statistical Analysis.

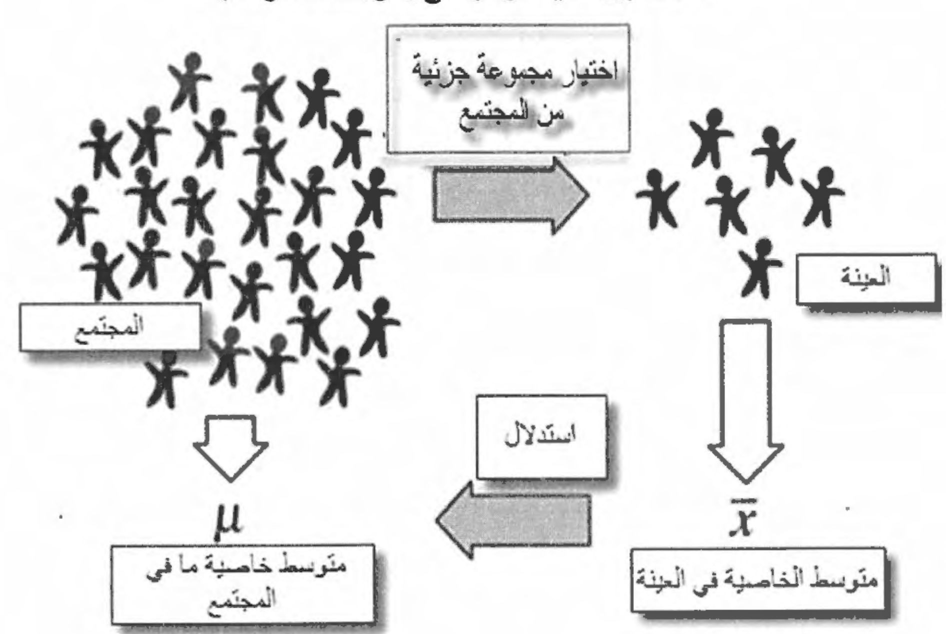
وهنا يكون لدينا متغيران تابعان أو أكثر، وصفر أو أكثر من المتغيرات المستقلة. وتشمل هذه الأساليب التحليل العاملي Factor Analysis، تحليل التباين للمتغيرات المتعددة MANOVA، ... إلخ.

الفصل الثالث
العينات الإحصائية Statistical Samples

العينة Sample:

تعرف على أنها ذلك الجزء من المجتمع الذي يتم الحصول عليه - عشوائياً أو غير عشوائي - لدراسة خصائص المجتمع والاستدلال على معالمه (انظر الشكل ٣-١)، ويتم اختيار العينة وفقاً لإحدى أساليب وطرق اختيار العينات.

شكل رقم (٣-١)
العلاقة بين العينة والمجتمع (من إعداد المؤلف)



لماذا العينات؟

إن تطبيقات استخدام العينات في الحياة العملية ضرورية جداً ولا يمكن الاستغناء عنها في شتى مجالات الحياة. ومن الأسباب التي تجعلنا نلجأ لأسلوب المعاينة عوضاً عن أسلوب الحصر أو المسح الشامل لكل وحدات المجتمع ما يلي:

١- تقليل الوقت والجهد والتكلفة المالية المطلوبة للدراسة.

٢- صعوبة حصر جميع مفردات أو وحدات المجتمع مثل أن يكون المجتمع غير قابل للعد مثل مخزون المياه، الأمطار، أو أن يكون المجتمع منتشراً في مساحات يصعب معها حصر جميع مفرداته مثل انتشار الأسماك في البحار، والحيوانات البرية في الصحاري والغابات ... إلخ.

٣- عدم إتلاف كل وحدات مجتمع الدراسة. كاختبار جودة نوع من المنتجات - مثل المعلبات الغذائية - والتأكد من مطابقته للمواصفات المطلوبة، فليس من الحكمة والمنطق اختبار كل وحدة منتجة وإتلافها للتأكد من أنها مطابقة للمواصفات.

٤- الرغبة في زيادة دقة القراءات والقياسات التي يتم الحصول عليها من وحدات أو مفردات الدراسة. فأسلوب الحصر الشامل، ولا سيما في حالة المجتمعات الكبيرة الحجم تحتاج إلى عدد أكبر من الأفراد الذين يقومون بجمع البيانات وترميزها وتفريغها، ومما لا شك فيه أن هؤلاء الأفراد ليسوا على نفس المستوى من التأهيل والخبرة بعملية جمع البيانات وأخذ القياسات وتفريغها، ومن ثم فإن احتمالية نسبة الوقوع في أخطاء جمع البيانات وتفريغها أعلى منه في حالة أسلوب العينات التي يكون فيها في الغالب فرصة استقطاب العدد المطلوب من المؤهلين في جمع وتسجيل وتفريغ البيانات متاح.

إحصاءة (مؤشر) العينة Sample Statistic:

إحدى خصائص العينة والتي يتم قياسها كمياً من خلال عينة مسحوبة من مجتمع الدراسة وتستخدم كمقدر لمعلمة المجتمع، كحساب المتوسط الحسابي $\bar{x}$ لمتغير ما من العينة واستخدام قيمته كتقدير للمتوسط الحسابي μ لذلك المتغير في المجتمع.

وحدة المعاينة Sampling Unit:

هي كل وحدة أو مجموعة من الوحدات يتم اختيارها أو سحبها من مجتمع الدراسة في كل مرحلة من مراحل المعاينة. فمثلاً في حالة العينة العشوائية العنقودية ذات المراحل المتعددة يتم سحب مجموعة عشوائية من العناقيد في المرحلة الأولى (مجموعة من المدارس الابتدائية في مدينة ما: المدرسة تمثل وحدة معاينة وهي تحوي مجموعة من الصفوف الدراسية) ثم يتم سحب مجموعة عشوائية من العناقيد في المرحلة الثانية (مجموعة من فصول الصف الخامس ابتدائي: الصف يمثل وحدة معاينة

وهو يحتـوي على مجموعة من الطلاب)، ثم في المرحلـة الثالثة والأخيرة يتم اختيار (مجموعـة مـن الطلاب من كل صف: الطالب يمثل وحـدة معاينة وهي الوحدة التي سيتم إجراء القياسـات عليها). ولأن الطالب هو الوحدة التي سيتم إجراء القياسات عليه أو جمع البيانات منه، فإنه يسمى وحدة التحليل أو القياس.

وحدة التحليل Analysis Unit أو وحدة القياس Measurement Unit:

وهي الوحدة التي يتم أخذ البيانات أو القياسات منها حسب متغيرات الدراسة. وقد تكون وحدة التحليل أو القياس هي نفسها وحدة المعاينة أو تختلف عنها. فمثلاً لنفرض أن هناك شـاحنة محملة بنوع معين من الفاكهة القادمة للمملكة العربية السعودية من إحدى الدول وأراد المختصون العاملون في الجمارك التأكد من سـلامتها وصلاحيتها للاسـتهلاك البشري. هنا يقوم المختص بسـحب عينة عشوائية من الصناديق. فإذا كان هـدف المختص فحص جميع حبـات الفاكهة الموجودة بكل صندوق في العينة فإن الصنـدوق في هذه الحالة يعتبـر وحدة معاينة ووحدة تحليل أو قياس. أما إذا اكتفى بفحص حبة فاكهة مـن كل صندوق فإن الصندوق في هذه الحالة يعتبر وحدة معاينة بينما حبة الفاكهة المفحوصة تعتبر وحدة تحليل أو قياس. وفي الدراسات الاجتماعية والإدارية والتسويقية يكون الأفراد عادة هم وحدات التحليل.

إطار المعاينة Sampling Frame:

ويسمى أحياناً بإطار المجتمع وهو عبارة عن المجموعة التي تحوي وحدات المعاينة. أو يعرف على أنه قائمة تحتوي على جميع وحدات مجتمع الدراسة، وهو يعد المصدر الأساسـي لاختيار عناصر العينـة، ويجب على الباحث تحديده وإعـداده إن لم يكن متوفراً ومعداً مسـبقاً. وينبغي أن يكون إطار المعاينة حديثاً وشـاملاً لجميع مفردات مجتمع الدراسة وجميع الفئات والطبقات. وتجدر الإشارة إلى أنه يصعب تحديد إطار المعاينـة في حالة المجتمعات غير المحدودة - كالأسـماك في البحار - وهذا يقود إلى ما يسـمى بالتحيز في اختيار وحدات الدراسة، أي عدم تكافؤ فرص وحدات مجتمع الدراسة في أن تكون ضمن العينة.

أنواع العينات الإحصائية Statistical Samples:

تقسـم العينات الإحصائية بناءً على أسـاليب اختيارها إلى قسمين رئيسيين هما: العينات الاحتمالية (العشوائية) والعينات غير الاحتمالية (غير العشوائية) وكل قسم من

تلك العينات يضم أنواعاً مختلفة (فهمي، ٢٠٠٥؛ عاروري، ٢٠١٣) و(Daniel, 2012). وسيتم هنا الاقتصار على تلك الأنواع من العينات الأكثر شيوعاً واستخداماً في الدراسات والبحوث.

أولاً - العينات الاحتمالية (العشوائية) Probability(Random) Samples:

وهي التي تقوم على مبدأ أن أي وحدة من وحدات المجتمع لها احتمال أو فرصة محددة ومعروفة لأن تكون ضمن عينة الدراسة. ومصطلح العشوائية هنا لا يعني «كيفما اتفق» وإنما يعني أن الفرصة متاحة لكل مفردة من مفردات المجتمع لأن تكون ضمن العينة بدون استثناء. إن العينات الاحتمالية تضبط نسبة التحيز في نتائج البحوث والدراسات، وتهدف إلى أن تكون العينة ممثلة تمثيلاً جيداً لمجتمع الدراسة ومن ثم تعميم نتائج الدراسة على المجتمع، كما أنه يمكننا تقدير دقة المؤشرات الإحصائية المحسوبة من العينة من خلال ما يعرف بالخطأ المعياري للمقدر.

ومن أشهر أنواعها وأكثرها استخداماً:

١- العينة العشوائية البسيطة Simple Random Sample:

وهذا النوع من العينات يتم استخدامه عندما تكون وحدات المجتمع متجانسة في متغير/ متغيرات الدراسة أو الظاهرة محل الدراسة. والمقصود بتجانس وحدات المجتمع هو أن تكون وحدات المجتمع متماثلة وغير مختلفة بشكل كبير بعضها عن بعض في الظاهرة محل الدراسة وغير متوزعة في مجموعات أو طبقات مختلفة. فمثلاً نجد أن العوائل السعودية التي تسكن في منطقة معينة من مدينة الرياض كالشمال تمثل مجتمعاً متجانساً في متوسط دخلها السنوي، في حين أن العوائل السعودية التي تسكن مناطق مختلفة من مدينة الرياض لا تمثل مجتمعاً متجانساً فمتوسط دخل العائلة السعودية السنوي يختلف باختلاف المنطقة (شمال، جنوب، شرق، غرب، وسط). وفي العينة العشوائية البسيطة تكون فرصة اختيار المفردة أو الوحدة من المجتمع مستقلة (أي لا تتأثر باختيار الوحدات الأخرى) ومتساوية لجميع مفردات المجتمع. ولاختيار وحدات العينة يتم ترقيم جميع وحدات المجتمع المحدد بإطار المعاينة، ومن ثم يتم الاختيار عشوائياً إما باستخدام برامج الحاسب الآلي أو باستخدام بعض الطرق التقليدية كجداول الأرقام العشوائية أو القصاصات الورقية والقبعة.

وتتميز العينة العشوائية البسيطة بالتالي:

- تعطي فرص متكافئة ومستقلة لمفردات المجتمع لأن تكون ضمن العينة.
- لا تتطلب معرفة مسبقة بخصائص المجتمع ولا تتقيد بترتيب معين أو نظام مقصود.
- أسهل في الفهم من حيث المفهوم وطريقة التنفيذ.
- لها ميل لأن تكون ممثلة لمجتمع الدراسة.
- تتفادى التحيز لاعتمادها إلى حد كبير على قانون الاحتمالات.

ومن عيوب العينة العشوائية البسيطة التالي:

- ينبغي أن تكون وحدات المجتمع متجانسة، وإلا فإن ذلك سيؤدي إلى عدم الدقة في النتائج.
- تحتاج إلى مجهود كبير ودقة لتحديد إطار للمعاينة في حالة عدم توفره مسبقاً.
- صعوبة استخدامها مع المجتمعات الإحصائية الكبيرة جداً، فتكون المعاينة مكلفة مادياً وتحتاج إلى وقت وجهد أكبر.

٢- العينة العشوائية المنتظمة Systematic Random Sample:

وهي أحد أنواع العينات العشوائية ويتم فيها اختيار وحدات العينة على فترات أو مسافات منتظمة من إطار المعاينة ويتم اختيار العينة العشوائية المنتظمة وفقاً لترتيب الخطوات التالية:

- تحديد إطار المعاينة ومن ثم ترتيب وحدات المعاينة وفقاً لخاصية معينة يحددها الباحث من ١ إلى N (حجم المجتمع).
- تحديد حجم العينة المرغوب فيه n.
- تحديد فترة الانتظام k من خلال $k = N/n$.
- اختيار الوحدة الأولى عشوائياً من الفترة من الوحدات التي تقع في الفترة ١ وk.
- أضف إلى العدد المختار قيمة k بشكل منتظم، واستمر في الاختيار حتى تصل إلى حجم العينة المطلوب n.

ومن مميزات العينة العشوائية المنتظمة ما يلي:

- من أسهل العينات العشوائية في التنفيذ، وأقلها جهداً وتكلفة.
- ليس هناك ضرورة لحصر جميع مفردات المجتمع وإنشاء إطار معاينة إذا كان هناك تمثيل مادي لوحدات المجتمع.

- تضمن انتشـــار مفـــردات المعاينة في المجتمـــع، وهذا ما لا يتم ضمانـــه في العينة العشوائية البسيطة، ومن ثم تقلل فرصة الارتباط الذاتي بين الوحدات.

وكما أن لها مميزات فإن لها عيوباً من أهمها:

- قد تكون العينة المختارة غير ممثلة لمجتمع الدراســـة، فقد يتم اختيار مفردات على أبعـــاد منتظمة يصـــادف أن يكونوا من طبقة معينة أو مـــن ذوي خصائص وصفات مميزة وغير متشابهة مع بقية المفردات في المجتمع.
- لا تحـــدث احتماليـــة فرصة الاختيار العشـــوائي إلا مرة واحدة، وهـــي عند اختيار المفردة الأولى فإن شرط الاستقلالية في الاختيار للوحدات الأخرى غير متوفر؛ لأن اختيار وحدات العينة يعتمد على ترتيب المفردة الأولى المختارة عشوائياً.
- المؤشـــرات الإحصائية كالتباين مثلاً أكثر تعقيداً في الحســـاب من العينة العشوائية البسيطة.

٣- العينة العشوائية الطبقية Stratified Random Sample:

في حالات عديدة تكون وحدات المجتمع غير متجانسة في الظاهرة محل الدراسة ولذا فإن المعاينة العشوائية البسيطة كثيراً ما تصبح غير ملائمة للاستخدام في هذه الحالة حيث إن اســـتخدامها يؤدي إلى الحصـــول على عينة تبتعد في خصائصها عن خصائـــص المجتمع مما يترتـــب عليه خطأ المعاينة. ولزيـــادة فرصة تمثيل خصائص المجتمع فى العينة، فإننا نلجأ إلى العينة العشـــوائية الطبقية. ففي العينة العشـــوائية الطبقية يتم تقســـيم المجتمع إلى طبقات على أســـاس متغير واحد أو أكثر مثل النوع (ذكر، أنثى)، المســـتوى التعليمي (متعلم، غير متعلم) ... إلخ، بحيث تكون الوحدات أو المفردات داخل كل طبقة متجانســـة ومن ثم يتم الاختيار عشوائياً من داخل كل طبقة. ويتم اختيار العينة العشوائية الطبقية بعدة أساليب منها:

- أســـلوب التوزيع المتساوي Equal Allocation: وفيه يكون أحجام العينات المسحوبة أو المختارة من الطبقات متساوياً.
- أســـلوب التوزيع النســـبي Proportional Allocation: وفيـــه تكون أحجام العينات المختارة من الطبقات حسب نسبتها في المجتمع.
- أسلوب التوزيع الأمثل Optimum Allocation: وفيه تكون أحجام العينات المختارة من كل طبقة حســـب عـــدة اعتبارات متمثلة فـــي أحجام الطبقـــات، التجانس بين الوحدات في كل طبقة، التكلفة.

ومن مميزات العينة العشوائية الطبقية ما يلي:

- درجة تمثيل العينة للمجتمع الأصلي ولطبقاته الفرعية الكبيرة والصغيرة عالياً.
- حجم خطأ المعاينة في العينة العشوائية الطبقية أقل منه في عينة عشوائية بسيطة لنفس الحجم.
- دقة النتائج الإحصائية وانخفاض نسبة الخطأ المعيارى للمقدرات، خاصة كلما كانت المجموعات أو الطبقات متجانسة داخلياً.

ومن عيوبها ما يلي:

- تتطلب من الباحث المعرفة بشكل جيد لمجتمع دراسته لتحديد الطبقات اللازمة.
- تتطلب من الباحث إجراءات أكثر ومجهوداً أكبر وتكلفة أعلى من العينة العشـوائية البسـيطة لسـحب عدد من العينات تبعاً لعدد مسـتويات المتغير أو المتغيرات الذي يتعامل معها .
- حساب المؤشرات الإحصائية كالمتوسط والتباين في العينة الطبقية أكثر تعقيداً منها في حالة العينة العشوائية البسيطة والمنتظمة.
- تتطلب وجود إطار معاينة لكل طبقة من الطبقات.

٤- العينة العشوائية العنقودية Cluster Random Sample:

في واقعنا العملي يوجد مواقف وحالات كثيرة يصعب معها استخدام أسلوب المعاينة العشـوائية البسـيطة أو المنتظمة أو حتى الطبقية. فقد يكون حجم مجتمع الدراسة كبيراً جداً ومنتشراً على نطاق جغرافي واسع، ومن ثم يصعب تحديد إطار معاينة له، بالإضافـة إلى الوقت والجهد والتكلفة العالية المطلوبـة، وصعوبة إدارة ومتابعة فريق عمل جمع البيانات. فإذا كان هناك تجانس بين وحدات الدراسـة في متغير الدراسة وكان مجتمع الدراسـة كبيراً ومنتشراً على نطاق جغرافي واسع يصعب معه استخدام العينة العشـوائية البسيطة أو المنتظمة، فإنه يتم تقسيم المجمع إلى مجموعات تسمى العناقيـد وفقاً لمعيار معين مثل المعيار الجغرافي، بحيث يراعى أن يكون التجانس بين العناقيد أكبر ما يمكن، ومن ثم يتم اختيار عينة عشـوائية من تلك العناقيد. والعينة العشوائية العنقودية تنقسم إلى عدة أقسام منها:

- العينة العشوائية العنقودية البسيطة أو ذات المرحلة الواحدة: وفيها يتم اختيار عينة من العناقيد ومن ثم يتم اختيار كل الوحدات دخل كل عنقود.

- العينـــة العشـــوائية العنقودية ذات المرحلتين: وفيها يتم اختيـــار عينة من العناقيد، وبدلاً من اختيار كل محتويات العناقيد يتم اختيار عينة عشـــوائية من المفردات من داخل كل عنقود تم اختياره في المرحلة الأولى.
- العينة العشوائية العنقودية المتعددة المراحل: وفيها يتم اختيار عينة من العناقيد في المرحلة الأولى وفقاً لمعيار معين وتسمى العناقيد في هذه المرحلة بالوحدات الأولية، ثم بعد ذلك يتم اختيار عينة من العناقيد وفقاً لمعيار آخر من داخل الوحدات الأولية وتســـمى العناقيد في هذه المرحلة بالوحدات الثانوية، وهكذا حتى نصل إلى المرحلة الأخيرة المتمثلة في سحب كل أو بعض وحدات المعاينة من العناقيد النهائية التي تم اختيارها.

ومن مزايا العينات العشوائية العنقودية ما يلي:

- تناسب المجتمعات الكبيرة المتناثرة التى تشغل حيزاً جغرافياً شاسعاً.
- لا تحتاج إلى إطار معاينة لكل وحدات المجتمع، وإنما تكتفي بإطار المعاينة للمجموعات المختارة.
- تتطلب جهداً أقل وتكلفة أقل من العينة العشوائية البسيطة والمنتظمة.

ومن عيوبها ما يلي:

- تتطلب خطوات كثيرة لاختيار وحدات الدراسة تبعاً لعدد المراحل.
- مستوى تمثيلها لمجتمع الدراسة منفخض.
- تتطلب حجم عينة أكبر مقارنة بالعينة البســيطة والمنتظمة إذا ما أراد الباحث رفع مستوى تمثيلها لمجتمع الدراسة وتقليص خطأ المعاينة.
- حســاب المؤشرات الإحصائية كالمتوســط والتباين في العينة العنقودية أكثر تعقيداً منها في حالة العينة العشوائية البسيطة والمنتظمة.

والفـــرق الجوهري بـــين العينة العشـــوائية العنقودية والطبقية هـــو أنه في العينة العشوائية العنقودية يتم تقسيم المجتمع إلى مجموعات أو عناقيد وفقاً لخاصية معينة كالتوزيع الجغرافي لوحدات المجتمع مع مراعاة أن يكون التجانس بين العناقيد أكبر ما يمكن، ومن ثم يتم اختيار عينة عشـــوائية من تلك العناقيد لدراستها. أما في المعاينة العشـــوائية الطبقية فإنه يتم تقســـيم المجتمع إلى مجموعات أو طبقات شـــاملة وغير متقاطعة حســـب خاصية أو أكثر مثل النوع (ذكر/أنثى)، المؤهل العلمي ... إلخ بحيث

تكـون الطبقات مختلفة فيما بينها والوحدات داخل كل طبقة متجانسـة، ومن ثم يتم اختيار عينة عشوائية من كل طبقة.

ثانياً - العينات غير الاحتمالية (غير العشوائية) Non Probability (Non Random) Samples:

هناك مواقف وظروف عديدة يكون معها استـخدام العينات العشوائية غير ملائم أو مستحيلاً. ومن تلك المواقف والظروف ما يلي:

- الحاجة إلى اتخاذ قرار سريع.
- الحاجة لدراسة وحدات أو مفردات محددة في المجتمع.
- حساسية وصعوبة الوصول إلى وحدات المجتمع، مثل مجتمعات المدمنين أو مروجي المخدرات.
- الموارد المالية والبشرية محدودة جداً.
- هدف البحث أو الدراسة استطلاعي.
- حساسية مجتمع الدراسة.
- الدراسة ليست ذات أهمية كبيرة وليس هناك حاجة للتعميم وإجراء استدلال إحصائي.
- صعوبة تحديد إطار للمعاينة لوحدات المجتمع.

فـي العينات غير العشـوائية لا يتم الاعتماد على مبدأ الاحتمالات أو العشـوائية في اختيار وحدات الدراسـة، بمعنى أن وحدات مجتمع الدراسة ليس لديها فرصة أو احتمـال محدد في أن تكون ضمن وحدات العينـة. وهذا يعنى أن الباحث يتحيز في اختياره لوحدات الدراسـة، ومن ثم فإن إمكانية التعميم مـن العينة إلى المجتمع غير ممكنة أو مقيدة بشـكل كبير جداً. وحيث إن العينـات غير الاحتمالية لا تعتمد على مبـدأ الاحتمالات في اختيار وحداتها فإنه لا يمكن معها تقدير خطأ المعاينة ولا يصح تقدير الخطأ المعياري للمؤشرات (المقدرات) الإحصائية واستخدام اختبار الفرضيات الإحصائيـة. وإنما يكتفي الباحث بوصف عينته واسـتخراج المؤشـرات الإحصائية المطلوبـة من العينة وعدم اسـتخدامها في التعميم إلا في حالات مقيدة جداً، سـيتم إيضاحهـا لاحقاً في هذا الفصل من الكتاب في جزئية بعنوان «إضاءة إحصائية حول استخدام أساليب الاستدلال الإحصائي في حالة العينات غير العشوائية».

هناك أنواع كثيرة للعينات غير العشـوائية ومن أكثر أنواعها اسـتخداماً وشـيوعاً ما يلي:

١- العينة الحصصية Quota Sample:

وهـــي أحد أنواع العينات غير العشـــوائية وهي تتطلب معرفة مســـبقة من الباحث لمجتمع الدراسة من حيث تكوين المجموعات والطبقات داخله. وفيها يتم تقسيم المجتمع إلى طبقات أو حصص Quota حسب خاصية أو معيار يحدده الباحث. وعملية اختيار وحدات الدراسة من كل مجموعة لا تتم بطريقة عشوائية ولكن تتم وفقاً لقناعة الباحث بـــأن وحدات كل مجموعـــة أو حصة Quatum في العينة تمثل المجموعة أو الحصة في المجتمع. وتعتبر العينة الحصصية من أفضل أنواع العينات غير العشوائية لأن الباحث يختار العينة وفقاً لخصائص أو معايير محددة مسبقاً لأفراد المجتمع. وتختلف العينة الحصصية عن العينة العشـــوائية الطبقية فـــي أن اختيار وحدات المعاينة من كل طبقة يكون عشوائياً في حالة العينة العشوائية الطبقية، أما في حالة العينة الحصصية فإنه يتم اختيار العينة من كل طبقة أو حصة بطريقة غير عشوائية.

٢- العينة الغرضية (أو الحكمية) Purposive (or Judgmental) Sample:

وفيهـــاً يتـــم اختيار وحدات أو مفردات الدراســـة بطريقة حرة غير عشـــوائية من مجتمـــع خاص لديه القدرة على توفير حاجة الباحث من البيانات والمعلومات المطلوبة وفقـــاً لغرض أو هدف محدد. ويتم اســـتخدام هذا النوع مـــن المعاينة عندما يتعامل الباحـــث مع حالات معينة يريد منهـــا بيانات ومعلومات خاصة حيـــث يختار الباحث أفضـــل المفردات أو الوحدات التي لديها القدرة أكثـــر من غيرها في توفير المعلومات وتحقيـــق الغرض أو الهدف الـــذي اختيرت من أجله. فمثلاً يمكن اســـتخدام العينة الغرضية لدراسة خصائص المديرين التنفيذيين الناجحين في قطاع الأعمال. وتتميز العينة الغرضية بأنها تضمن اختيار وحدات مناسبة للدراسة، كما أنها توفر كثيراً من الجهد والمال والوقت.

٣- العينة الملائمة أو المتاحة Convenient Sample:

وهـــذا النوع يعتبر من أضعف أنواع العينات غير العشـــوائية نظراً لارتفاع نســـبة التحيـــز في الاختيار فيهـــا وانخفاض درجة تمثيلها لمجتمع الدراســـة، حيث يتم فيها اختيار مفردات الدراســـة نتيجة لأنها متاحة فقط وليس لأي عامل آخر. ومثال على العينة المتاحة استطلاع رأي المتدربين الموجودين في صالة المطعم بمعهد الإدارة في أحد الأيام التدريبية، وذلك عن جودة الحقائب التدريبية. فهنا المتدربون الذين تم استطلاع رأيهم لم يتم اختيارهم عشوائياً، وإنما كانوا متاحين لحظة تنفيذ الاستطلاع.

٤- عينة كرة الثلج Snowball Sample:

وفيها يتصل الباحث بمفردة أو أكثر من مفردات المجتمع وهذه المفردة أو المفردات تقوده إلى مفردة أخرى، وهكذا حتى لا يستطيع الباحث الوصول إلى مفردات إضافية أو يصل الباحث إلى العدد الذي يراه مناسباً لدراسته. وهذا النوع من العينات يتطلب من الباحث القدرة على إقناع من يتصل معهم في مجتمع الدراسة في إرشاده إلى مفردات أخرى في المجتمع. وغالباً تستخدم عينة كرة الثلج فى حالة عدم توفر قائمة بكل أفراد المجتمع الأصلي. ومن أمثلة استخدامها دراسة مدمني المخدرات والعصابات الإجرامية.

إضاءة إحصائية حول استخدام أساليب الاستدلال الإحصائي في حالة العينات غير العشوائية:

في حالة العينات غير العشوائية، هل يمكن استخدام الإحصاء الاستدلالي inferential Statistics ومن ثم تعميم نتائج العينة على المجتمع الذي أخذت منه؟

لقد سبقت الإشارة إلى أنه يجب أن تكون العينة المسحوبة من مجتمع الدراسة ممثلة للمجتمع تمثيلاً جيداً للوصول إلى استنتاجات صحيحة حوله. فإذا تمكن الباحث من تحديد إطار المعاينة في دراسته والذي يعني أن لدى الباحث تصوراً واضحاً عن تركيبة مجتمع الدراسة فإن هذا سيمكنه من اختيار تصميم المعاينة المناسب حسب أهداف وتساؤلات وطبيعة مجتمع الدراسة، ومن ثم اختيار مفردات العينة بطريقة عشوائية (احتمالية) تمكن كل مفردة من مفردات المجتمع أن يكون لديها فرصة (احتمال) محددة لأن تكون ضمن مفردات العينة دون تحيز. وهذا يفسر لنا سبب تفضيل العشوائية عند اختيار العينات، فالمعاينة العشوائية ضرورية - ولكنها غير كافية - لأن تكون العينة ممثلة لمجتمع الدراسة (Glass & Hopkins, 1995). إن البعض يرى أنه يمكن استخدام الإحصاء الاستدلالي في حالة العينات غير العشوائية ولكن ينبغي أن يكون الباحث يقظاً وحذراً جداً عند التعميم من العينة إلى المجتمع (Huck, 2012, 100). حيث يشترط أن تكون العينة تمثل خصائص المجتمع المستهدف بالتعميم تمثيلاً جيداً، وهذا قد يتحقق في العينة الحكمية مثلاً، ولكنه قد لا يتحقق أو يصعب تحقيقه في العينة الملائمة. وهنا نود الإشارة إلى أنه يجب على الباحث عند تحليل ووصف بيانات عينة دراسته - سواء كانت العينة عشوائية أو غير عشوائية - أن يفصل ويتحرى الدقة في ذلك حسب أهداف وتساؤلات الدراسة؛ لأن ذلك يعطي

الباحث نفسه والقارئ تصوراً عاماً عن خصائص المجتمع المستهدف بتعميم النتائج عليه. وفي المقابل ومن وجهة نظر النظرية الإحصائية هناك من يتحفظ على تعميم نتائج العينة غير العشوائية - أياً كان نوعها - على المجتمع خاصة في حال استخدام الاختبارات الإحصائية (اختبارات الفرضيات) أو فترات الثقة، لكونها تعتمد على إحصائيات اختبار Test Statistics، وهذه الإحصائيات متغيرات عشوائية لها توزيعات احتمالية تسمى توزيعات المعاينة Sampling Distribution تعتمد على العينات العشوائية (الاحتمالية)، ومبدأ نظرية الاحتمالات، وهذا ما لا يتوافر في العينات غير العشوائية (غير الاحتمالية). كما أن هناك جانباً آخر ينبغي أخذه بعين الاعتبار وهو أن استخدام العينات غير العشوائية يؤدي إلى أو يزيد من احتمالية الوقوع في مشكلة التحيز Biasedness (والذي يعني اختيار - بقصد أو بدون قصد - عناصر معينة من مجتمع الدراسة تتمتع بخصائص معينة تختلف عن بقية عناصر المجتمع ولا تمثل كل أو معظم خصائص مجتمع الدراسة)، وهذا يؤدي إلى استقراءات واستنتاجات غير صحيحة لمعالم المجتمع، ومن ثم يفقد الدراسات البحثية مصداقيتها وموثوقيتها ولا يمكن الاعتماد عليها. وقد شدد Lajer (2007) على عدم استخدام الاختبارات الإحصائية في العينات غير العشوائية. كما أن Dale (2006) وGorard (2006) لهما رأي حول هذا الموضوع حيث ذكرا بأن استخدام اختبارات المعنوية غير مناسب في حالة العينات غير العشوائية لأن الاختبار الإحصائي يتعامل مع الخطأ من النوع الأول أو ما يعرف بمستوى المعنوية - سيتم الحديث عنه لاحقاً في الفصل الخامس - والذي بدوره يعتمد على العينة العشوائية.

خطوات إجراء المعاينة Sampling Procedures:

- ينبغي أن تتم عملية التخطيط والتنفيذ للمعاينة وفق الترتيب التالي:
- تحديد مجتمع الدراسة.
- تحديد إطار المعاينة إذا أمكن.
- تحديد أسلوب/أساليب المعاينة لاختيار العينة من المجتمع.
- تحديد حجم العينة أو العينات.
- تحديد وحدات المعاينة و/أو وحدات التحليل.
- تنفيذ المعاينة وجمع البيانات من وحدات التحليل.

مصادر الأخطاء في المعاينة Error Sources in Sampling:

إن البيانات والمؤشرات التي يتم الحصول عليها من العينة لا يمكن أن تمثل مجتمع الدراسة تمثيلاً تاماً وإنما تعبر عنه بدرجة معينة من الدقة تزيد أو تنقص وفقاً للمنهجية والإستراتيجية المتبعة في تنفيذ المعاينة. فمثلاً عند حساب متوسط العينة لتقدير متوسط المجتمع فإن متوسط العينة قد لا يطابق متوسط المجتمع على الأرجح بل يزيد أو ينقص عنه بمقدار معين، وهذا يرجع إلى أن أي عينة لا تخلو من أخطاء تتراوح في حجمها حسب درجة الدقة والاهتمام في تنفيذ الدراسة. وعليه فإن خطأ المعاينة يعرف على أنه الفرق بين قيمة المقدر (كالمتوسط أو النسبة مثلاً) الذي يتم حسابه من العينة والقيمة الحقيقية للمقدر في المجتمع (المعلمة). ويمكن تقسيم الأخطاء في المعاينة حسب مصادرها إلى قسمين رئيسيين هما:

أخطاء المعاينة العشوائية Random Sampling Error:

نتيجة لأخذ جزء من عناصر المجتمع (عينة) وليس كل عناصره فإنه من المتوقع أن تكون المؤشرات الإحصائية المحسوبة من العينة تختلف عن تلك التي في المجتمع حتى وإن تم اختيار تلك العينة بعناية فائقة. وهذا النوع من الخطأ لا يمكن تقليل تأثيره إلا من خلال زيادة حجم العينة (Madsen, 2011).

الأخطاء المنتظمة (أو غير العشوائية) Systematic Errors:

وتسمى أيضاً أخطاء التحيز Bias، وهذا النوع من الأخطاء يحدث بسبب عوامل معينة يمكن للباحث تحديدها وإزالة أثرها بمعالجة تلك العوامل ومن تلك العوامل المسببة للأخطاء المنتظمة في المعاينة ما يلي:

١- أخطاء إدارية وتشمل أخطاء في أسلوب المعاينة، وفي تنفيذ المعاينة، أخطاء متعلقة بعدم تحديد إطار المعاينة بالشكل السليم، أخطاء في طريقة المقابلات، أخطاء في صياغة أسئلة الاستبانات أو استطلاع الرأي، أخطاء في معالجة البيانات كترميزها وإدخالها في الحاسب الآلي وتحليلها.

٢- أخطاء تتعلق بأفراد العينة أو المستجيبين وتشمل عدم استجابة أفراد الدراسة في الإدلاء بالبيانات المطلوبة إما لرفضهم أو لعدم التمكن من الوصول إليهم. كما تشمل عدم صحة البيانات المتحصل عليها من المستجيبين بسبب عدم الاهتمام بالدراسة أو لعدم فهمها، أو أخطاء في القياس ... إلخ.

وتأسيساً على ذلك فإنه ينبغي ملاحظة ما يلي:

- أن الأخطاء المنتظمة غير العشوائية (التحيز) لا تقل بزيادة حجم العينة، وهذا ومن ثـم يؤثر على دقة بيانات ومؤشـرات العينة. وهناك مفهـوم خاطئ يقع فيه الكثير من الباحثـين وهو اعتقادهم أنه بمجرد زيادة حجـم العينة في ظل وجود الأخطاء المنتظمة فإن دقة نتائج العينة تزيد وهذا غير صحيح.
- وفي المقابل فإن الخطأ العشـوائي يقل بزيـادة حجم العينة ومن ثم تزيد دقة نتائج العينـة. أي أنـه إذا تم اختيار أسـلوب المعينة الصحيح وتم تنفيـذه وفق المنهجية العلمية الصحيحة وتم معالجة البيانات بالشـكل السليم فلا شك في أن زيادة حجم العينة سيؤدي إلى زيادة الدقة في النتائج ولاسيما في ظل وجود انخفاض في درجة التجانس بين وحدات الدراسة.

إضاءات إحصائية حول العينات:

- إن معظـم المكاتب والمراكز الإحصائية ومراكز اسـتطلاع الرأي وقياس الاتجاهات تفضل أساليب المعاينة العشـوائية وتوصي بالاهتمام بتركيبة المجتمع الديموغرافية وعدم إغفالها أثناء المعاينة.
- في حال كون مجتمع الدراسـة معقداً في تركيبتـه فإنه ليس بالضرورة وجود عينة مثاليـة وحيـدة ممثلة لذلـك المجتمع، وإنما قد يتم اسـتخدام مزيج مـن العينات الاحتماليـة وغير الاحتماليـة، وذلك للوصول إلى أفضل عينـة ممثلة تحوي كل أو معظم صفات المجتمع.
- إذن يمكن القول بأن العينة الجيدة - سـواء كانت عشـوائية أو غير عشوائية - هي تلك التي تعكس خصائص مجتمع الدراسة بشكل كبير.
- ينبغي وصف بيانات عينة الدراسـة وصفاً جيداً ودقيقاً، سواء كانت العينة عشوائية أم غير عشوائية، لأن ذلك يساعد في وصف واستقراء معالم مجتمع الدراسة الذي سـحبت منه العينة، ومن ثم يمكننا تعميم نتائج العينة على المجتمع الذي يمتلك تلك المعالم أو الخصائص. بمعنى آخر، إن مؤشـرات (إحصاءات) العينة تعطينا تصوراً لمعالم وملامح المجتمع الذي يمكننا تعميم النتائج عليه.
- أن هناك الكثير من الأبحاث التي تفيد بأنه تم استخدام المعاينة العشوائية وهي في الحقيقة لم تسـتخدمها نهائياً بل اسـتخدمت المعاينة غير العشوائية أو أنها أساءت استخدامها مما يؤدي إلى نتائج غير صحيحة للدراسة.

تحديد حجم العينة Sample Size Determination:

إن موضوع حجم العينة من الأمور الهامة والحساسة جداً في الإحصاء والدراسات والبحوث. وهناك تساؤل يتم طرحه دائماً حول حجم العينة المناسب لإجراء أي تحليل إحصائي معين أو أي دراسة تعتمد على جمع البيانات وتحليلها. نظرياً وعملياً ينبغي أن لا تكون العينة صغيرة جداً أو كبيرة جداً، بل يجب أن يكون حجمها كافياً بحيث يفي بالغرض من الدراسة وهو استقراء ما يريده الباحث عن المجتمع بشكل صحيح ودقيق من خلال عينة دراسته، والاستقرار Stability في نتائج الدراسة . فلا ينبغي أن تكون العينة صغيرة جداً مما يزيد من احتمال عدم احتواء العينة على كل أو معظم صفات وخصائص مجتمع الدراسة، ومن ثم لا يمكن الاعتماد على نتائج تحليل العينة في استقراء صفات وخصائص ذلك المجتمع بشكل دقيق، وأن لا تكون العينة كبيرة جداً فتكون العينة في هذه الحالة غير عملية (فقدان الهدف من المعاينة) كونها ستكون غالباً مكلفة جداً وتحتاج إلى وقت وجهد كبيرين، بالإضافة إلى بعض الاعتبارات الفنية الأخرى (الوصول إلى استنتاجات غير صحيحة) المتعلقة بأساليب التحليل الإحصائي. وفي أدبيات حجم العينات وتأثيرها على نتائج الاختبارات الإحصائية ذكر - على سبيل المثال لا الحصر - Tabacknick & Fidell (2007) أن العينات الصغيرة جداً قد تؤدي إلى استنتاج أنه لا توجد علاقة ذات دلالة إحصائية بين متغيرات الدراسة مع وجودها في الواقع، كما أن العينات الكبيرة جداً في الحجم قد تؤدي إلى استنتاج أن هناك علاقة ذات دلالة إحصائية بين متغيرات الدراسة مع عدم وجودها في الواقع. وكمثال على ذلك فإن قيمة معامل الارتباط الخطي لبيرسون Pearson's Correlation Coefficient (r) بين متغيرين كميين قد تكون صغيرة جداً (مثلا ٠,١٠ = r أي أن العلاقة بين المتغيرين ضعيفة جداً أو معدومة)، ومع ذلك وباستخدام الاختبار الإحصائي للعلاقة الخطية بين المتغيرين يمكن الوصول إلى أن هناك علاقة خطية معنوية بين المتغيرين في حال كان حجم العينة كبير، وهذا موضع استفهام سببه كبر حجم العينة عن الحد المطلوب.

وفي الواقع العملي فإن هناك طريقتين أو منهجين مختلفين في عملية تحديد حجم العينة وهما:

١- الاعتماد على الخبرات والتجارب السابقة للخبراء والإحصائيين Rule of Thumb.

٢- الاعتماد على القواعد والنظريات الاحتمالية في الإحصاء.

وفيما يلي سيتم تناول تحديد حجم العينة في حالة المعاينة العشوائية والمعاينة غير العشوائية.

أولاً - حجم العينة في حالة المعاينة العشوائية:

إن عملية تحديد حجم العينة ولاسيما في الدراسات والبحوث التي يوجد بها العديد من المتغيرات وتحتاج لتنفيذ العديد من أساليب التحليل الإحصائي المختلفة، يعتبر من المهام الصعبة التي يواجهها الباحث والمختص، وذلك لعدم وجود قانون أو صيغة رياضية موحدة تصلح لتحديد حجم العينة أياً كانت الدراسة أو البحث.

إن تحديد حجم العينة المطلوب والمرغوب فيه يعتمد على مجموعة من العوامل التي ينبغي أخذها بعين الاعتبار، وهي كالتالي:

١- أهداف البحث وفرضياته.

٢- حجم مجتمع الدراسة وانتشاره.

٣- مدى تباين أو تجانس بيانات متغير/متغيرات الدراسة في المجتمع.

٤- هامش الخطأ أو درجة الدقة في التقدير (Margin of Error/Accuracy of estimation) وهي تعكس درجة الدقة في تقدير معالم المجتمع ويتم عادة استخدام القيم من ١٪ إلى ٥٪ وتعتبر ٥٪ أكثرها استخداماً.

٥- مستوى الثقة (Confidence level) لتقدير معالم المجتمع، فمثلاً قد يريد باحث - ومن خلال بيانات عينة الدراسة - أن يكون على ثقة مقدارها ٩٥٪ بأن متوسط الدخل السنوي لموظفي شركة قطاع خاص معينة يتراوح بين قيمتين محددتين، ويتم عادة استخدام مستويات الثقة ٩٠٪، ٩٥٪ أو ٩٩٪ وتعتبر ٩٥٪ أكثرها استخداماً.

٦- نوع المتغيرات المستخدمة في الدراسة، كمية Quantitative أم نوعية Categorical.

- في حالة المتغيرات الكمية فإنه يتم عادة تقدير المتوسطات، ومن ثم يمكن استخدام الصيغة التالية لتقدير حجم العينة المطلوبة لتقدير متوسط المتغير الكمي:

$$n = \left(z_{a/2} * \frac{\sigma}{e}\right)^2$$

حيث:

n : الحد الأدنى لحجم العينة المطلوب.

$z_{a/2}$: القيمـــة الحرجة المرتبطة بمســتوى ثقة مقدارهـــا (1-α) ٪ ويتم الحصول عليها من جدول التوزيع الطبيعي المعياري، α هي مستوى المعنوية أو مستوى الدلالة الإحصائية.

σ: الانحـــراف المعياري للمتغيـــر في المجتمع، ويمكن تقديره من العينة إذا لم يكن معلوماً.

e : هامـــش الخطأ (أو دقة التقدير) وهو الحد الأعلى لمقدار الخطأ المســموح به في التقدير.

وعادة يأخذ إحدى القيم من ١٪ إلى ٥٪ حسب الدراسة أو البحث.

وهذا الحجم هو الحجم المســتهدف للعينة ويتم اعتباره واستخدامه إذا كان حجم المجتمـــع كبير (Infinite Population) مقارنة بحجم العينة، وقد جرت العادة على اســتخدام القيمـــة ٥٪ للحكم على ما إذا كان حجم العينـــة صغيراً مقارنة بحجم المجتمع (Daniel, 2012). فإذا كانت نسبة حجم العينة إلى حجم المجتمع (وهذا ما يعرف بكسر المعاينة ($f=\frac{n}{N}$) أصغر من ٥٪ فإن المجتمع يعدُّ كبير الحجم مقارنة بالعينـــة. أما إذا كان حجم المجتمـــع صغيراً أو محـــدوداً (Finite Population) مقارنة بحجم العينة، أي إذا كان كسر المعاينة أكبر من ٥٪، فإنه يتم تصحيح حجم العينة بحيث يصبح حجم العينة المستهدف هو (أبو شعر، ١٩٩٧):

$$n_1=\frac{n}{1+f}=\frac{n}{1+\left(\frac{n}{N}\right)}$$

حيث:

n : يتم حســـابها من الصيغة الســـابقة وN ترمز لحجـــم المجتمع وn_1 تمثل حجم العينة النهائي.

- أما إذا كان المتغير محل الدراســـة نوعياً، فإنه يتم عادة تقدير النســـب، ومن ثم يمكن اســـتخدام الصيغة الرياضية التاليـــة لتقدير حجم العينة المطلوب لتقدير نسبة المتغير النوعي:

$$n=\left(\frac{z_{a/2}*\sqrt{p(1-p)}}{e}\right)^2$$

حيث سبق تعريف n ، $z_{a/2}$ و e. أما p فإنها ترمز لنسبة الظاهرة في المجتمع كنسبة السعوديين العاملين في شركة معينة بالنسبة للعاملين فيها. غالباً ما تكون هذه القيمة مجهولة، وقد جرت العادة على إعطائها القيمة ٠,٥ وهي القيمة التي تجعل حجم العينة أكبر ما يمكن، وهذا يتحقق عندما يكون تباين الظاهرة محل الدراسة عند قيمته العظمى.

وهذا الحجم يتم اعتباره واستخدامه إذا كان حجم المجتمع كبيراً جداً مقارنة بحجم العينة. أما إذا كان حجم المجتمع صغيراً أو محدوداً مقارنة بحجم العينة، فإنه يتم تصحيح حجم العينة بحيث يتم استخدام حجم العينة التالي:

$$n_1 = \frac{n}{1+f} = \frac{n}{1+\left(\frac{n}{N}\right)}$$

حيث:

n : يتم حسابها من الصيغة السابقة و N ترمز لحجم المجتمع و n_1 تمثل حجم العينة النهائي.

وينبغي التنويه إلى أن تلك الصيغ الرياضية السابقة لتحديد حجم العينة لمتغير واحد فقط ما هي إلا وسيلة مساعدة في تحديد حجم العينة النهائي المطلوب لإنجاز دراسة أو بحث تشتمل على مجموعة من المتغيرات المتنوعة.

٧- أساليب التحليل الإحصائي المستخدمة كتقدير مؤشرات المجتمع، مقارنة مجموعات، مقاييس الارتباط بين المتغيرات، نماذج التنبؤ ... إلخ.

٨- عدد الطبقات و/أو المجموعات في المجتمع.

٩- نوع تصميم المعاينة المستخدم كالعينة العشوائية البسيطة، العينة العشوائية الطبقية، العينة العشوائية العنقودية، ... إلخ.

١٠- عدد المتغيرات في الدراسة.

١١- درجة أهمية الدراسة ودقتها المطلوبة. فكلما ارتفعت درجة أهمية الدراسة وكانت دقتها المطلوبة كبيرة، كان حجم العينة المطلوب أكبر.

١٢- نوع البحث أو الدراسة (دراسة كمية أو نوعية، أو مزيج من الكمي والنوعي).

١٣- معدل الاستجابة Response Rate المتوقع من المبحوثين أو عناصر الدراسة.

١٤- التكلفة والوقت والجهد اللازم للحصول على العينة.

كل تلـــك العوامل أو بعضها ينبغي أخذها بعين الاعتبـــار عند تحديد الحد الأدنى المطلوب لحجم العينة حسب نوع وطبيعة وظروف الدراسة أو البحث.

إضاءات إحصائية حول حجم العينات العشوائية (١):

- في الدراسات والأبحاث التي تتناول الأساليب الإحصائية المتعددة المتغيرات كالتحليل العاملي Factor Analysis، تحليل الانحدار المتعدد Multiple Regression والتحليل التمييـــزي Discernment Analysis ... إلخ، فقـــد تباينت الآراء حول الحد الأدنى للعينة، فهناك دراسات تقترح ٥ وحدات معاينة (أفراد أو جماعات أو غير ذلك حسب الهدف من الدراسة) لكل متغير من متغيرات الدراسة ، والبعض اقترح ١٠ وحدات، وآخرون يرون استخدام ١٥ أو ٢٠ وحدة لكل متغير من متغيرات الدراسة. وأكثر تلك الآراء قبولاً هو أن يتم استخدام ١٠ وحدات معاينة لكل متغير من متغيرات الدراسة كحد أدنـــى (Bartlett, Kotrlik &Higgins., 2001). فمثلاً لو أراد الباحث إجراء تحليـــل انحدار متغيـــر تابع معين على ١٠ متغيرات مســـتقلة فإنه يلزمه ١٠٠ وحدة معاينـــة على الأقـــل (حاصل ضرب عدد المتغيرات المســـتقلة (١٠) في عدد وحدات المعاينـــة (١٠وحدات) لكل متغير مســـتقل)، وقد ذكر بارتليت وآخـــرون أن التحليل العاملي لا ينبغي إجراؤه بما لا يقل عن ١٠٠ وحدة.
- هناك البعض من الباحثين يســـتخدم جداول تقديـــر حجم العينة التي تم تصميمها وتقديمها من قبل باحثين آخرين، فعلى ســـبيل المثال قام كلٌّ من كريجســـي ومورقان Krejcie & Morgan (1970) بتقـــديم جـــدول لتحديد الحـــد الأدنى لحجم العينة المطلـــوب للمتغيـــرات النوعيـــة Categorical variables، وقد اســـتخدما الصيغة الرياضية التالية في تصميم ذلك الجدول.

$$n = \frac{x^2_{1,1\text{-}a} \, NP\,(1\text{-}p)}{e^2(N\text{-}1) + x^2_{1,1\text{-}a} \, P\,(1\text{-}p)}$$

حيث:

n: حجم العينة المطلوب.

$x^2_{1,1\text{-}a}$: قيمة مربع كاي الجدولية بدرجة حرية تساوي الواحد وبمستوى ثقة محددة $(1\text{-}a)\%$.

N وP: سبق تعريفهما.

e: هامش الخطأ (أو دقة التقدير) وهو الحد الأعلى لمقدار الخطأ المسموح به في التقدير. وعادة يأخذ إحدى القيم من ١٪ إلى ٥٪ حسب الدراسة أو البحث.

جدول رقم (٣-١)
الحد الأدنى لحجم العينة (n) من المجتمع ذو الحجم (N) حسب هامش الخطأ (e) ومستوى الثقة (CL) باستخدام قانون كريجسي ومورقان لتحديد حجم العينة (تم استخدام برنامج إكسل Excel لإنشاء هذا الجدول)

	e=0.05			e=0.01		
	CL=0.90	CL=0.95	CL=0.99	CL=0.90	CL=0.95	CL=0.99
N	*n*	*n*	*n*	*n*	*n*	*n*
١٠	١٠	١٠	١٠	١٠	١٠	١٠
٢٠	١٩	١٩	١٩	٢٠	٢٠	٢٠
٣٠	٢٧	٢٨	٢٩	٣٠	٣٠	٣٠
٥٠	٤٢	٤٤	٤٧	٥٠	٥٠	٥٠
٧٥	٥٩	٦٣	٦٧	٧٤	٧٤	٧٥
١٠٠	٧٣	٨٠	٨٧	٩٩	٩٩	٩٩
١٥٠	٩٧	١٠٨	١٢٢	١٤٧	١٤٨	١٤٩
٢٠٠	١١٥	١٣٢	١٥٤	١٩٤	١٩٦	١٩٨
٢٥٠	١٣٠	١٥٢	١٨٢	٢٤١	٢٤٤	٢٤٦
٣٠٠	١٤٣	١٦٩	٢٠٧	٢٨٧	٢٩١	٢٩٥
٤٠٠	١٦٢	١٩٦	٢٥٠	٣٧٨	٣٨٤	٣٩١
٥٠٠	١٧٦	٢١٧	٢٨٥	٤٦٦	٤٧٥	٤٨٥
٦٠٠	١٨٧	٢٣٤	٣١٥	٥٥١	٥٦٥	٥٧٩
٧٠٠	١٩٥	٢٤٨	٣٤١	٦٣٤	٦٥٣	٦٧٢
٨٠٠	٢٠٢	٢٦٠	٣٦٣	٧١٥	٧٣٩	٧٦٣
٩٠٠	٢٠٨	٢٦٩	٣٨٢	٧٩٤	٨٢٣	٨٥٤
١٠٠٠	٢١٣	٢٧٨	٣٩٩	٨٧١	٩٠٦	٩٤٣

تابع - جدول رقم (٣-١).

	e=0.05			e=0.01		
	CL=0.90	CL=0.95	CL=0.99	CL=0.90	CL=0.95	CL=0.99
N	*n*	*n*	*n*	*n*	*n*	*n*
١٢٠٠	٢٢١	٢٩١	٤٢٧	١٠١٩	١٠٦٧	١١١٩
١٥٠٠	٢٢٩	٣٠٦	٤٦٠	١٢٢٨	١٢٩٧	١٣٧٦
٢٠٠٠	٢٣٨	٣٢٢	٤٩٨	١٥٤٤	١٦٥٥	١٧٨٥
٢٥٠٠	٢٤٤	٣٣٣	٥٢٤	١٨٢٦	١٩٨٤	٢١٧٣
٣٥٠٠	٢٥١	٣٤٦	٥٥٨	٢٣٠٧	٢٥٦٥	٢٨٩٠
٥٠٠٠	٢٥٧	٣٥٧	٥٨٦	٢٨٧٥	٣٢٨٨	٣٨٤٢
٧٥٠٠	٢٦١	٣٦٥	٦١٠	٣٥٥٧	٤٢١١	٥١٦٥
١٠٠٠٠	٢٦٣	٣٧٠	٦٢٢	٤٠٣٥	٤٨٩٩	٦٢٣٩
٢٥٠٠٠	٢٦٨	٣٧٨	٦٤٦	٥٢٢٤	٦٩٣٩	٩٩٧٢
٥٠٠٠٠	٢٦٩	٣٨١	٦٥٥	٥٩٥٨	٨٠٥٦	١٢٤٥٥
٧٥٠٠٠	٢٧٠	٣٨٢	٦٥٨	٦٢٠٤	٨٥١٤	١٣٥٨٣
١٠٠٠٠٠	٢٧٠	٣٨٣	٦٥٩	٦٣٣٥	٨٧٦٢	١٤٢٢٧
٢٥٠٠٠٠	٢٧٠	٣٨٤	٦٦٢	٦٥٨٦	٩٢٤٨	١٥٥٥٥
٥٠٠٠٠٠	٢٧٠	٣٨٤	٦٦٣	٦٦٧٤	٩٤٢٣	١٦٠٥٥
١٠٠٠٠٠٠	٢٧٠	٣٨٤	٦٦٣	٦٧١٨	٩٥١٢	١٦٣١٧
٢٥٠٠٠٠٠	٢٧١	٣٨٤	٦٦٣	٦٧٤٦	٩٥٦٧	١٦٤٧٨
١٠٠٠٠٠٠٠	٢٧١	٣٨٤	٦٦٣	٦٧٥٩	٩٥٩٤	١٦٥٦٠
١٠٠٠٠٠٠٠٠	٢٧١	٣٨٤	٦٦٣	٦٧٦٣	٩٦٠٣	١٦٥٨٤

- يمكن أيضاً استخدام طريقة كوهن (1988) Cohen والتي تعتمد على أسلوب تحليل القوة الإحصائية Statistical Power Analysis لاختيار الحد الأدنى للعينة. وقد أشار شوان (2006) Chuan إلى أن طريقة كوهن هي واحدة من أكثر الطرق المستخدمة في تحديد حجم العينات في مجال علم السلوك. وهي تعتمد في تحديدها على عدة خصائص إحصائية تتضمن حجم التأثير effect size، مستوى المعنوية α، وقوة الاختبار الإحصائي المرغوبة في اكتشاف الفروق بين المجموعات أو العلاقات بين المتغيرات. وهذه الطريقة تحتاج إلى إلمام جيد بمفهومها وإلى برامج حاسوبية متخصصة لحساب حجم العينة. ويمكن حساب حجم العينة باستخدام طريقة كوهن من خلال البرامج الحاسوبية المتخصصة مثل R، SAS، وSPSS.
- أن تكون العينة ذات حجم كبير لا يضمن أن تكون ممثلة للمجتمع تمثيلاً جيداً، وإنما يعتمد تمثيل العينة للمجتمع بشكل جيد على الإستراتيجية Sampling Strategy التي يتم بها اختيار عناصر العينة.
- كما أن حجم العينة يعتمد على مدى تجانس أو تباين المجتمع في خصائص الدراسة. فكلما كان المجتمع متجانساً كانت العينة أصغر في الحجم، والعكس صحيح.

خطوات عملية تحديد حجم العينة المناسب في حالة العينات العشوائية البسيطة والمنتظمة:

بعد التفصيل السابق وكل الاعتبارات المتعلقة بتحديد حجم العينة اللازم لدراسة أو بحث ما، فإنه قد يكون من المفيد أن يتم تقديم الخطوات العملية التي يتم استخدامها في إجراء الدراسات المسحية survey research - والتي تعتبر من أكثر أنواع الدراسات استخداماً من قبل الباحثين والدارسين - وفي مراكز استطلاع الرأي وقياس الاتجاهات لتحديد الحد الأدنى للعينات.

١- تحديد حجم مجتمع الدراسة المراد سحب العينة / العينات منه. وينبغي معرفة أن حجم معظم مجتمعات الدراسة (المجتمع الإحصائي) يصعب أو لا يمكن تحديده بشكل دقيق، وإنما يتم تقديره. لذا يلزم الاجتهاد والاهتمام بشكل كبير بتعريف وتحديد حجم مجتمع الدراسة لكي لا يفقد جزءاً هاماً من أجزاء المجتمع.

٢- تحديد هامش الخطأ Margin of Error المتوقع في تقدير المجتمع من العينة، حيث في التطبيق العملي يمكن إعطاؤه أحد القيم من ١٪ الى ٥٪ حسب ظروف الدراسة أو البحث.

٣- تحديد مســـتوى الثقة Confidence Level في تقديـــر المجتمع من العينة، وعملياً يعطـــى مســـتوى الثقة أحـــد القيم مـــن ٩٠٪ ، ٩٥٪ أو ٩٩٪ وتعتبـــر ٩٥٪ أكثرها استخداماً وشيوعاً.

٤- يتـــم حســـاب (أو تحديد) حجم العينة المطلـــوب اللازم لدراســـته (وليكن n) بناء على المعطيات الســـابق ذكرها في الخطوات من ١-٣، وذلك من خلال اســـتخدام إحدى الصيغ الرياضية أو الأساليب السابق ذكرها لتحديد حجم العينة. الجدير بالذكر أن جدول كريجســـي ومورقان شـــائع الاســـتخدام بين كثير من الباحثين، (Chuan, 2006). ويرجح المؤلف اســـتخدام صيغة كريجســـي ومورقان حيث إنها تعطي حجم عينة أكبر من غيرها من الطرق الأخرى لحساب حجم العينة في حالة المتغيرات الكمية أو النوعية تحت نفس الظروف المحددة مثل فترة الثقة ومســـتوى المعنوية ألفا وهامش الخطأ في التقدير.

٥- يتـــم تقديـــر معدل الاســـتجابة Response Rate ويرمز له بالرمـــز r، حيث نتوقع دائماً أن يكون هناك نســـبة معينة من عناصر عينة الدراسة لا يتم الحصول على بياناتهـــا، أو أن البيانات التي تم الحصول عليها من تلك العناصر جزئية وناقصة ولا تصلح للدراسة ومن ثم يتم استبعادها. وهذه النسبة قد تكون كبيرة أو صغيرة حسب الدراسة، نوع المبحوثين (عناصر الدراسة)، توقيت الدراسة، نوعية البحث، علاقة الباحث بالمبحوثين، أســـئلة الدراســـة وغيرها من المتغيـــرات التي قد تؤثر بشـــكل مباشر أو غير مباشر في معدل الاســـتجابة. فإذا كان الباحث لديه خبرة أو معرفة مسبقة بمعدل الاستجابة في المجال الذي يبحث فيه فإنه يتم استخدام تلك النسبة في حساب حجم العينة النهائي اللازم لدراسته. أما إذا لم يكن لديه خبرة أو معرفة مسبقة بمعدل عدم الاستجابة فإنه - وحسب ما توصي به المراكز المتخصصة في مجال استطلاع الرأي وقياس الاتجاهات - ولضمان الجانب الآمن فيما يتعلق بحجم عينة الدراسة المطلوب، يتم تقدير معدل الاستجابة بـ ٥٠٪ وهي القيمة التي تقع في منتصف الفترة (٠ - ١٠٠) ٪ لمعدل الاستجابة.

٦- ومن هنا يكون حجم العينة النهائي (المستجيبين الذين سيتم دعوتهم للمشاركة في الدراسة أو البحث) وليكن n^* هو:

$$n^* = \frac{n}{r}$$

إضاءات إحصائية حول حجم العينات العشوائية (٢):

- الطريقة الســـابقة لتحديد حجم العينة تفترض أســـلوب العينة العشوائية البسيطة والتي تفترض أن مجتمع الدراسة متجانس في متغيرات الدراسة، أي أنها لا تفترض أن عناصر المجتمع مقسم إلى طبقات حسب متغيرات الدراسة، وهذا ما ينبغي أخذه بعين الاعتبار.

ولذا إذا كان الباحث سيســـتخدم أسلوب المعاينة العشـــوائية الطبقية فإن هناك حالتـــين لكيفية التعامل مع تحديد حجـــم العينة، وهذا يعتمد على درجة الدقة التي يرغبها الباحث في نتائجه (Bankier, 1988).

الحالــة الأولــى: إذا تم تحديد هامش الخطـــأ للمجتمع بأكمله بغض النظر عن طبقاتـــه فإنه في هذه الحالة يقوم بتطبيق الخطوات الســـابقة من ١ الى ٦ لتحديد حجم العينة الكلي ومن ثم يقوم بتوزيعه على الطبقات إما بالتساوي أو حسب حجم الطبقة في المجتمع الكلي أو بطريقة التوزيع الأمثل.

الحالة الثانية: إذا تم تحديد هامش الخطأ لكل طبقة من طبقات المجتمع، فإنه في هذه الحالة يقوم بتطبيق الخطوات الســـابقة من ١ الى ٦ لحســـاب حجم العينة لكل طبقة، ومن ثم يكون الحجم الكلي لعينة دراســـته هو مجموع أحجام العينات في تلك الطبقات، وســـينتج في هذه الحالة حجم عينة كلي أكبر من الحالة الأولى مما يقلل من نسبة خطأ المعاينة.

- في حالة العينات العشوائية العنقودية، ينبغي تحديد عدد العناقيد التي سيتم اختيارها عشوائياً من عناقيد مجتمع الدراسة. وهناك أساليب إحصائية أكثر تقدماً لتحديد عـــدد العناقيد المراد اختيارها في العينة وتعتمد في حســـابها علـــى معرفة بتباين العناقيد في المجتمع، والتي تكون غالباً غير معلومة وتحتاج إلى تقدير بطريقة معينة كاستخدام عينة استطلاعية، كما تعتمد على متوسط حجم العنقود في المجتمع ونوع المعلمة المراد تقديرها كالمتوســـط الحسابي والنسبة، وكذلك مقدار الخطأ المسموح به في تقدير معلمة المجتمع. وللاطلاع على تلك الصيغ الرياضية يمكن الرجوع إلى الكتب الإحصائية المتخصصة في المعاينة ومنها (عاروري، ٢٠١٢).

وينبغي التنويه إلى أن هناك عاملين أساسيين يؤثران في تحديد عدد العناقيد المـــراد اختيارها تتمثـــل في عامل الوقت والمـــال والجهد وعامـــل الدقة أو درجة تمثيل العينة لمجتمع الدراســـة. فإذا كان الوقت والتكلفة المالية والجهد التي سيتم

استـــتنفاذها للوصول إلى تلك العناقيد كبيراً فإنه يتـــم عادة تقليص عدد العناقيد إلـــى عدد أقل من المطلوب، ومن ثم ســـيكون احتمالية درجـــة تمثيل العينة لمجتمع الدراسة أقل.

في الدراسات البحثية والتي تحتوي عادة على متغيرات بحثية كثيرة كمية ونوعية وأســـاليب تحليل إحصائي مختلفة يتم تحديد حجم العينة العنقودية الكلي المطلوب للدراسة كالتالي:

١- يتـــم تحديد حجم العينة الكلي n^* باســـتخدام الخطوات الســـابقة من ١ إلى ٦ التي تســـتخدم في حالة العينة العشوائية البســـيطة، ومن ثم يتم حساب حجم العينـــة العشـــوائية العنقودية الكلـــي n^*_c والذي ينبغي أن لا يقل عن n^* بســـبب ما يســـمى بتأثير تصميـــم المعاينـــة Sampling Design Effect (deff) والذي عرفه Kish (1992) على أنه نســـبة تباين المقدر (إحصاءة أو مؤشر العينة مثل المتوســـط الحسابي أو النســـبة مثلاً) Estimator باســـتخدام أي نوع من أنواع المعاينة العشـــوائية إلى التباين في المقدر باستخدام المعاينة العشوائية البسيطة، أي أن:

$$deff = \frac{\text{تباين المقدر باستخدام أي نوع من أنواع المعاينة العشوائية}}{\text{تباين المقدر باستخدام العينة العشوائية البسيطة}}$$

إن قيمة تأثير تصميم المعاينة سيكون مساوياً للواحد الصحيح ($deff = 1$) في حالة العينة العشوائية البسيطة والمنتظمة وأقل من الواحد الصحيح ($deff < 1$) في حالة العينة العشـــوائية الطبقيـــة وأكبر من الواحد الصحيح ($deff > 1$) في حالة العينة العشوائية العنقودية. إن تحديد حجم تأثير تصميم المعاينة لدراسة معينة يمكن من خلال الدراسات السابقة المماثلة أو من خلال عينة استطلاعية. أما إذا تعذر إمكانية تحديد أثر تصميم المعاينة العشوائية الطبقية أو العنقودية فإنـــه يمكن اختيار $deff = 1$ للعينة الطبقية و$deff \geq 2$ (أي ٢ أو أكثر) في حالة العينة العنقودية (Statistics Canada, 2010).

٢- يتـــم تحديـــد حجم العينة العشـــوائية العنقوديـــة الكلي n^*_c من خـــلال المعادلة التالية:

$$n^*_c = n^* \times deff$$

٣- بعد تحديد حجم العينة الكلي n^*_c ، يتم تقدير متوسـط حجم العينة الذي سيتم اختيـاره من كل عنقود في العينـة (أي $\bar{n}_c$) ومن ثم يتم تحديد عدد العناقيد m التي سيتم اختيارها عشوائياً من مجتمع الدراسة كالتالي:

$$m = n^*_c \div \bar{n}_c$$

مثـال: نفترض أن لدينا مجتمعاً مكوناً من ٥٠ عنقوداً، وأن حجم العينة الكلي - باستخدام الخطوات من ١-٦ المراد سحبها، ٥٠٠ وحدة معاينة (n^* = ٥٠٠)، وأنه تم تقدير قيمة تأثير التصميم من خلال دراسات سابقة أو من خلال عينة استطلاعية بـ ٢ (أي أن deff = 2) والذي بناء عليه سـيكون حجم العينة العشـوائية العنقودية الكلي يساوي:

$$n^*_c = n^* \times deff = 500 \times 2 = 1000$$

وأن متوسـط حجم العينة المراد سـحبها من كل عنقود هـو ١٠٠ وحدة معاينة (أي أن ١٠٠ = $\bar{n}_c$) ومـن ثـم فـإن عـدد العناقيـد التـي ينبغـي اختيارهـا في العينة يساوي:

$$m = n^*_c \div \bar{n}_c = 1000 \div 100 = 10$$

أي أنه سـيتم اختيار ١٠ عناقيد عشوائياً ومن ثم سيتم اختيار ١٠٠ وحدة تحليل عشـوائياً من كل عنقود من تلك العناقيد المختـارة، ومن ثم نحصل في النهاية على عينة عشوائية بحجم كلي يساوي ١٠٠٠ وحدة.

- من الأخطاء الشائعة المتعلقة بحجم العينات والتي يرتكبها كثير من الباحثين، أن يتم اختيار حجم عينة الدراسة مساوياً (٣٨٤) وحدة، وذلك في حالة المجتمعات الكبيرة الحجم أو ما تسـمى «بالمجتمعات غير المحدودة» إحصائيـاً Infinite Population وذلك بالاعتماد على جدول كريجسـي ومورقان أو بعض الصيغ الرياضية الخاصة بتحديـد حجم العينة. هذا قـد يكون صحيحاً في حالات معينة ولكنه وبلا شـك لا يصلح لكل الحالات وفي كل الظروف. وذلك لأن الرقم (٣٨٤) مبني على مجموعة من الافتراضات والتي منها أن يكون توزيع البيانات في مجتمع الدراسـة طبيعياً أو قريبـاً من الطبيعي، أن هامش الخطأ في التقدير مسـاوٍ لـ ٥٪، وفترة الثقة لتقدير معالم المجتمع ٩٥٪. والسؤال، هل سيبقى الحد الأدنى لحجم العينة المطلوب يساوي (٣٨٤) لـو أراد الباحث أن يكون هامش الخطأ في التقدير يسـاوي ١٪ وفترة الثقة

٩٥٪ مثـــلاً؟ طبعاً الإجابـــة وبالتأكيد: لا. وإليك التوضيـــح التالي الذي يبين مدى التغيـــر الكبير في حجم العينة المطلوب اختيارها من مجتمع حجمه مثلاً ١٠٠٠٠٠٠ (مليون) وحدة، وذلك بمجرد تغيير هامش الخطأ و/أو مستوى الثقة.

جدول رقم (٣-٢)

مقدار التغير في الحد الأدنى لحجم العينة باستخدام طريقة كريجسي ومورقان لمجتمع كبير الحجم (١٠٠٠٠٠ فأعلى) بناء على تغيير مقدار هامش الخطأ في التقدير ومستوى الثقة

هامش الخطأ (e)	مستوى الثقة	الحد الادنى لحجم العينة (n)
٥٪	٩٥٪	٣٨٤
٥٪	٩٩٪	٦٥٥
١٪	٩٥٪	٩٥١٣
١٪	٩٩٪	١٦١٢٠

إن الجـــدول (٣-٢) أعلاه يوضح أن حجـــم العينة المطلوب دالة في هامش الخطأ ومســـتوى الثقة فبمجرد تغيير هامش الخطأ من ٥٪ إلى ١٪ عند مســـتوى الثقة ٩٥٪ أصبح الحد الأدنى لحجم العينة العشوائية المطلوب ٩٥١٣ وحدة بدلاً من ٣٨٤ وحدة. وعليه فإنه ينبغي على الباحث أن يهتم جيداً بمدى ملاءمة حجم العينة لدراســـته وما هـــي الافتراضات والمبررات وظروف وطبيعة البحث التـــي بناء عليها تم اختيار حجم العينة، وأن يوضح ذلك في بحثه ودراســـته حتى تزداد درجة الثقة والدقة والمصداقية في نتائج دراسته.

دليل عملي لتحديد حجم العينة العشوائية المطلوب للمتغيرات النوعية:

إن الجدول (٣-٣) يلخص الخطوات العملية التي يمكن للباحث أن يستعين بها في عملية تحديد حجم العينة العشوائية البسيطة والمنتظمة، الطبقية، والعنقودية.

جدول رقم (٣-٣)

دليل عملي لتحديد حجم العينة العشوائية المطلوب للمتغيرات النوعية

نوع المعاينة العشوائية	الخطوة	الإجراء
	١	حدد حجم مجتمع الدراسة N.
	٢	حدد هامش الخطأ e والذي يتراوح بين ١٪ و٥٪
	٣	حدد مستوى الثقة $(1-a)$ ٪ وعادة يتم استخدام إحدى القيم ٩٠٪، ٩٥٪ أو ٩٩٪.
	٤	قدر نسبة ظهور الظاهرة أو المتغير الرئيسيي في مجتمع الدراسة P، وإذا لم يكن لدى الباحث معرفة بذلك فيمكنه إعطاء P القيمة ٠,٥ وهذه القيمة تجعل حجم العينة أكبر ما يمكن.
	٥	أوجد حجم n الأولية من المعادلة $n=\frac{x^2_{1,1-a}\,NP\,(1-p)}{e^2(N-1)\,x^2_{1,1-a}\,P\,(1-p)}$
	٦	قدر معدل الاستجابة المتوقع r. وإذا لم يكن لديك مستند نظري أو عملي عن القيمة المتوقعة لذلك فإنه يمكن جعل $r = 50\%$.
	٧	حدد قيمة تأثير تصميم المعاينة المستخدمة في الدراسة deff. - في حالة المعاينة العشوائية البسيطة والمعاينة العشوائية المنتظمة تكون $deff = 1$. - أيضاً في حالة المعاينة العشوائية الطبقية اجعل $deff = 1$. - في حالة المعاينة العشوائية العنقودية اجعل $deff \geq 1$. أشار Daniel (2012) إلى أن القيمة الافتراضية لـ deff في العينة العنقودية يساوي ٢.
	٨	حجم العينة النهائي n^* الذي سيتم اختياره في الدراسة حسب نوع العينة العشوائية هو:
العينة العشوائية البسيطة أو المنتظمة		$n^* = \frac{n}{r} \times deff = \frac{n}{r} \times 1 = \frac{n}{r}$
العينة العشوائية الطبقية		$n^* = \frac{n}{r} \times deff = \frac{n}{r} \times 1 = \frac{n}{r}$
العينة العشوائية العنقودية		$n^* = \frac{n}{r} \times deff = \frac{n}{r} \times 1 = \frac{n}{r} \times 2$ أي أنه يفضل أن يكون حجم العينة العشوائية العنقودية ضعف العينة العشوائية البسيطة على الأقل.

تطبيق عملي تفاعلي لتحديد حجم العينة العشوائية البسيطة المناسب:

تطبيق عملي تفاعلي (٣-١):

لنفرض أن لدينا منظمة معينة عدد موظفيها ١٠٠٠٠ موظف، ونريد سحب عينة عشوائية بسيطة منها لتقدير نسبة رضاهم عن الحوافز المادية والمعنوية التي تقدمها تلك المنظمة لمنسوبيها. ما الحد الأدنى لحجم العينة لتقدير نسبة رضا العاملين بتلك المنظمة بمستوى ثقة مقدارها ٩٥٪ وبهامش خطأ في التقدير لا يتجاوز ٥٪؟

الحل:

١- حجم مجتمع الدراسة N = 10000.

٢- هامش الخطأ في التقدير e = 0.05.

٣- مستوى الثقة = ٠,٩٥.

باستخدام جدول كريجسي ومورقان وتحت الظروف السابق تحديدها في الخطوات ١-٣ فإن حجم العينة اللازم هو n = 370 مستجيب. (نلاحظ أن نسبة حجم العينة إلى حجم المجتمع أقل من ٥٪ لذلك فإن حجم المجتمع يعتبر كبيراً، وعليه فإنه لن يتم تعديل حجم العينة).

٤- لنفرض أنه ليس لدينا خبرة سابقة أو توقع محدد بمعدل الاستجابة، لذا سنجعله يساوي ٥٠٪ (أي r = 0.50).

٥- ومن ثم سيكون حجم العينة النهائي (n^*) هو:

$$n^* = \frac{n}{r} = \frac{370}{0.50} = 740$$

أي أنه ينبغي على الباحث دعوة (اختيار) ٧٤٠ موظفاً في عينته للمشاركة في دراسته وذلك تحسباً لأن يكون ٥٠٪ من المستجيبين لن يستجيب أو سيفقد من العينة.

ثانياً - حجم العينة في حالة المعاينة غير العشوائية:

في نظرية الاحتمالات لا يمكننا استخدام الأساليب الإحصائية المعتمدة على التوزيعات الاحتمالية في تحديد حجم العينات غير العشوائية (غير الاحتمالية). كما أن أسلوب المعاينة غير العشوائية قد تكون الأنسب في مواقف عديدة لعدة اعتبارات تتمثل في الوقت المطلوب لتنفيذ الدراسة والتكلفة واعتبارات أخلاقية وتصميم الدراسة والهدف

منها وغير ذلك. وقد تناولت الأدبيات في هذا المجال الإرشـــادات والقواعد والقوانين العامة غير الإحصائية Rules of Thumb المستندة إلى آراء الخبراء والمتخصصين في مجـــال العلوم الإحصائية في تحديد حجم العينات غير العشـــوائية. والجدول (٣-٤) يعرض الحجم التقريبي المتعارف عليه بين الباحثين للعينة غير العشوائية حسب تصميم أو منهج البحث أو الدراسة، (Daniel, 2012, 243).

جدول رقم (٣-٤)
حجم العينة غير العشوائية حسب نوع الدراسة البحثية

تصميم (منهج) البحث أو الدراسة	حجم العينة (عدد مفردات العينة أو المشاركين)
دراسة حالة Case Study	٣ إلى ٥
بحوث الظواهر Phenomenological Research	٦ إلى ١٠
النظرية المتجذرة Grounded Theory	١٥ إلى ٣٠
الإثنوجرافية (الأعراق والأجناس) Ethnographic Research	٣٥ إلى ٥٠
المجموعات البؤرية Focus Group	٣ إلـى ١٢ مجموعة حسـب نوع المشاركين، ولكل مجموعة ٦ إلى ١٢ مشارك.
البحوث التجريبية Experimental Research	١٥ إلى ٣٠ لكل مجموعة
الدراسة المسحية ذات الموضوع الواحد على المستوى المجتمعي أو الوطني Survey single-topic community or national	٤٠٠ إلى ٢٥٠٠
الدراسة المسحية ذات المواضيع المتعددة على المستوى الوطني Survey multiple-topic national Study	١٠٠٠٠ إلى ١٥٠٠٠
البحوث الاستكشافية، الدراسات الاستطلاعية، الاختبار القبلي أو المبدئي Exploratory Research, Pilot Study, pretest	٢٠ إلى ١٥٠
البحوث الارتباطية Correlational Research	٣٠
تحليل المجموعات الفرعية الرئيسية Analysis of major subgroup	١٠٠
تحليـل المجموعات الفرعية الثانويـة أو الصغيرة Analysis of minor subgroup	٣٠
الدراسة التسويقية أو اختبار منتج Marketing research or product testing	٢٠٠ إلى ٢٥٠٠
مجتمع حجمه أكبر من ٤٠٠	٢٠٠ إلى ١٥٠٠

الفصل الرابع
أساليب التحليل الإحصائي الوصفية
Descriptive Statistical Analysis

مقدمة:

عبارة عن مجموعة من الأدوات والأساليب الرياضية تهدف إلى استكشاف ومعالجة وتنظيم وتلخيص البيانات للوصول إلى معلومات ذات قيمة مفيدة. فالإحصاء الوصفي يسلط الضوء على طبيعة الظاهرة محل الدراسة وسلوكيات توزيع بياناتها وتمركزها وتباينها وعلاقة متغيراتها وأبعادها ببعضها البعض وعلاقتها أيضاً بظواهر أخرى. كما أنه يساعد في اكتشاف القيم الشاذة أو المتطرفة أو المفقودة في البيانات واختيار أسلوب الاستدلال الإحصائي الأنسب. وتتمثل أساليب التحليل الوصفي في التالي :(Blueman, 2008; Antonius, 2013)

١- جداول التوزيع التكرارية Frequency Distribution Tables.

٢- الرسومات الإحصائية Statistical Graphs.

٣- مقاييس النزعة المركزية (المتوسطات) Central Tendency Measures ومن وأكثرها استخداماً وشيوعاً المتوسط الحسابي Mean، الوسيط Median، المنوال Mode.

٤- مقاييس التشتت Dispersion Measures ومن أكثرها استخداماً وشيوعاً المدى Range، المدى الربيعي Inter-Quartile Range، التباين Variance، الانحراف المعياري Standard Deviation، معامل الاختلاف Coefficient of Variation.

٥- مقاييس شكل التوزيع Shape of Distribution Measures وتتمثل في الالتواء Skewness والتفرطح Kurtosis.

٦- مقاييس الموضع النسبي Relative Position Measures:

وهذه المقاييس لا تستخدم بكثرة في الدراسات والأبحاث، وتشمل:

- الربيعيات Quartiles.
- العشيريات Deciles.
- المئينيات Percentiles.
- الدرجات المعيارية Standard Scores مثل درجات Z وT المعيارية.

٧- مقاييـــس العلاقة بين المتغيرات Relationship Measures ومن أشـــهرها وأكثرها اســـتخداماً معامل ارتباط بيرســـون Pearson's Correlation Coeffiecient (r) ومعامل ارتباط ســـبيرمان Spearman's Correlation Coeffient(Rho)، وسيتم تناولها بالتفصيل في الفصل الثامن.

ويعتمـــد الاختيار من بـــين تلك الأدوات على هدف الدراســـة ونوعيـــة المتغيرات ومســـتوى قياســـها. والجدول (٤-١) يوضح توزيع الأســـاليب الإحصائية الوصفية حســـب نوع المتغيرات ومستويات قياسها. وأســـاليب الإحصاء الوصفي أساسية في الدراسات والأبحاث ولا يمكن للباحث أو المهتم الاستغناء عنها. فعندما يقوم الباحث بجمـــع بيانات متغيرات الدراســـة تكون في صورة «بيانات خـــام» Raw data متناثرة يصعب على الباحث أو المســـتفيد الاســـتفادة منها في وضعها الخام بالشـــكل المفيد، ولاســـيما عندما يكون حجم البيانات كبير. لذا فإن عليه اســـتخدام أساليب التحليل الإحصائي الوصفيـــة لتنظيم بياناته بطرق علمية معينة وعملية تســـهل عليه وصف الظاهرة محل الدراســـة ومتغيراتها والعلاقة بينها والوصول إلى اســـتنتاجات مفيدة ذات معنى وقيمة.

جدول رقم (٤-١)

توزيع الأساليب الإحصائية الوصفية حسب نوع الأسلوب ونوع المتغيرات ومستوى قياسها

	نوع المتغيرات ومستوى قياسها			
	متغيرات نوعية		متغيرات كمية	
الأسلوب الإحصائي الوصفي	**اسمي**	**رتبي**	**فتري**	**نسبي**
الجداول التكرارية	الجداول التكرارية غير المبوبة (النوعية)		الجداول التكرارية المبوبة	
الرسومات الإحصائية	الأعمدة البيانية الدائرة البيانية		الخط البياني أو السلسلة البيانية المدرج التكراري المنحنى التكراري رسم الصندوق الشكل الانتشاري	

تابع - جدول رقم (٤-١).

	نوع المتغيرات ومستوى قياسها			
	متغيرات نوعية		متغيرات كمية	
الأسلوب الإحصائي الوصفي	اسمي	رتبي	فتري	نسبي
مقاييس النزعة المركزية	المنوال	الوسيط المنوال	المتوسط الحسابي الوسيط المنوال المتوسط الهندسي	
مقاييس التشتت	نسبة الاختلاف	معامل الاختلاف	المدى التباين والانحراف المعياري المدى الربيعي معامل الاختلاف	
مقاييس شكل التوزيع			الالتواء التفرطح	
مقاييس الموضع			الربيعيات العشيريات المئينيات الدرجات المعيارية	
مقاييس الارتباط	كرامير	سبيرمان كندال تاو جاما	معامل ارتباط بيرسون	

جداول التوزيع التكراري Frequency Distribution Tables:

الجـــداول التكراريـــة عبارة عن تنظيم وفرز للبيانات الخـــام في جداول مكونة من فئـــات Category or Classes وتكرارات Frequencies. ومن الأســـباب التي تجعلنا نستخدم الجداول التكرارية ما يلي:

- تنظيم وتلخيص وعرض البيانات الخام في صورة يجعلها أكثر وضوحاً وذات فائدة ومعنى.
- تمكن القارئ والمستفيد من التعرف على توزيع البيانات.
- تمكن الباحث والمحلل من عمل رسومات بيانية محددة كالمدرج التكراري أو الأعمدة البيانية مثلاً.
- تمكن الباحث والقارئ من عمل المقارنات بين مجموعات البيانات والتعرف على العلاقات بين المتغيرات.

وتقسم الجداول التكرارية حسب نوع البيانات إلى قسمين هما:

١- جداول التوزيع التكراري غير المبوبة Ungrouped Frequency Distribution Tables:

وهي خاصة بالمتغيرات النوعية ذات المقياس الاسمي أو الرتبي مثل متغير الحالة الاجتماعية، التخصص، المرتبة الوظيفية أو البيانات الكمية التي لها فئات قليلة مثل عدد أفراد الأسرة.

٢- جداول التوزيع التكراري المبوبة Grouped Frequency Distribution Tables:

وهي خاصة بتوزيع وتنظيم بيانات المتغيرات الكمية مثل الدخل، الطول، الوزن، ضغط الدم ... إلخ.

وجداول التوزيع التكراري قد تكون:

- بسيطة تمثل التوزيع التكراري لمتغير واحد فقط.
- أو متقاطعة أو مزدوجة (مركبة) تمثل التوزيع التكراري لمتغيرين أو أكثر في آن واحد.

كيفية بناء جداول التوزيع التكراري باستخدام برنامج SPSS من خلال تطبيقات عملية تفاعلية:

سيتم توضيح كيفية تنفيذ وبناء الجداول التكرارية للبيانات المبوبة وغير المبوبة باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٤-١):

البيانات في الجدول (٤-٢) تمثل استجابة عينة مكونة من ٢١ فرداً على أسئلة الاستبانة تم اختيارها عشوائياً من مجتمع الدراسة (العمر، الجنس، عدد السيارات المملوكة، نوع السكن، المؤهل العلمي، عدد الأطفال، الحالة الاجتماعية، الحالة الاقتصادية، الدخل الشهري بالريال السعودي):

جدول رقم (٤-٢)
البيانات الخام لمتغيرات العمر، الجنس، الطول، الوزن، عدد السيارات المملوكة، المؤهل العلمي، عدد الأطفال، الحالة الاجتماعية، الحالة الاقتصادية، الدخل الشهري

الدخل الشهري	الحالة الاقتصادية	الحالة الاجتماعية	عدد الأطفال	المؤهل العلمي	نوع السكن	عدد السيارات	الجنس	العمر
٩٦٣٦	متوسطة	مطلق	٥	أقل من ثانوي	ملك	٥	ذكر	٣٧
١٢٦٢٣	جيدة	متزوج	١	جامعي	ملك	٢	أنثى	٥٤
٨٩٩٦	متوسطة	مطلق	١	ثانوي	ملك	٥	ذكر	٦٥
٨٢٨١	متوسطة	أعزب	٦	ثانوي	ملك	٥	أنثى	٢٩
٧٠٩٠	سيئة	متزوج	١٠	ثانوي	ملك	٤	ذكر	٣٥
٦٠٠٢	متوسطة	مطلق	١	جامعي	ملك	٣	أنثى	٣٦
٥٢٠٨	سيئة	متزوج	٩	ثانوي	إيجار	٥	ذكر	٤٣
٥٢٤٥	متوسطة	متزوج	٢	ثانوي	ملك	٥	ذكر	٢٥
١١٤٠٦	جيدة	متزوج	٥	جامعي	إيجار	١	ذكر	٥٥
١٠٤٢٣	جيدة	متزوج	٤	جامعي	إيجار	٤	أنثى	٢٢
١١٥٩١	ممتازة	أرمل	٨	فوق الجامعي	إيجار	١	ذكر	٤٢
٥٧٤٤	سيئة	أرمل	٧	ثانوي	ملك	٥	أنثى	٣٢
١٤٣٠٢	ممتازة	متزوج	٤	فوق الجامعي	إيجار	٢	أنثى	٣٩
١٢٩١٨	ممتازة	أرمل	٢	جامعي	إيجار	٥	ذكر	٣٤
١٠٢١٤	ممتازة	مطلق	٠	جامعي	إيجار	٤	أنثى	٣١
١١٣٠٠	جيدة	متزوج	٥	فوق الجامعي	إيجار	٠	أنثى	٤٧
٨٥١٠	متوسطة	متزوج	٨	جامعي	إيجار	٤	ذكر	٦٢
٩٦٢٩	جيدة	مطلق	٥	جامعي	إيجار	٥	أنثى	٤٣
١٢٦٤٠	ممتازة	متزوج	٣	فوق الجامعي	ملك	٤	ذكر	٦٥
١١٩١١	ممتازة	متزوج	٤	جامعي	إيجار	٥	ذكر	٢٥
١٥١٧٥	ممتازة	أعزب	٦	فوق الجامعي	إيجار	٤	ذكر	٢٥

المصدر (بتصرف): (فهمي، ٢٠٠٥).

المطلوب:

إدخال تلك البيانات في برنامج SPSS ومن ثم إيجاد ما يلي:

١- جدول التوزيع التكراري لمتغير الحالة الاقتصادية.

٢- جدول التوزيع التكراري لمتغير العمر بعد تحويله إلى فئات طولها ١٠ سنوات لكل فئة.

٣- جدول التوزيع التكراري لمتغيري الحالة الاقتصادية والعمر.

خطوات الحل:

١- قـم بإدخال المتغيرات والبيانات الموضحة في الجدول (٤-١) أعلاه إلى SPSS من خـلال نافذتـي Variable View وData View ومن ثم قم بحفظ الملف، لتحصل على الملف التالي:

شكل رقم (٤-١)

المتغيرات والبيانات الواردة في التطبيق العملي التفاعلي (٤-١) بعد إدخالها في برنامج SPSS

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

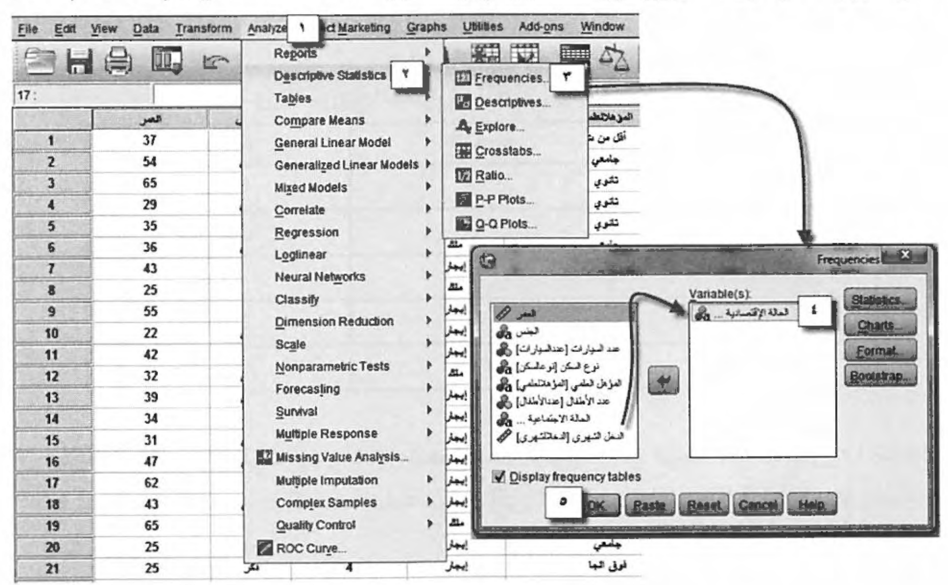
30 :

	العمر	الجنس	عدد السيارات	نوع المسكن	المؤهل العلمي	عدد الأطفال	الحالة الاجتماعية	الحالة الاقتصادية	الدخل الشهري
1	37	ذكر	5	ملك	أقل من ث	5	مطلق	متوسطة	9636
2	54	أنثى	2	ملك	جامعي	1	متزوج	جيدة	12623
3	65	ذكر	5	ملك	ثانوي	1	مطلق	متوسطة	8996
4	29	أنثى	5	ملك	ثانوي	6	أعزب	متوسطة	8281
5	35	ذكر	4	ملك	ثانوي	10	متزوج	سيئة	7090
6	36	أنثى	3	ملك	جامعي	1	مطلق	متوسطة	6002
7	43	ذكر	5	إيجار	ثانوي	9	متزوج	سيئة	5208
8	25	ذكر	5	ملك	ثانوي	2	متزوج	متوسطة	5245
9	55	ذكر	1	إيجار	جامعي	5	متزوج	جيدة	11406
10	22	أنثى	4	إيجار	جامعي	4	متزوج	جيدة	10433
11	42	ذكر	1	إيجار	فوق الجا	8	أرمل	ممتازة	11591
12	32	أنثى	5	ملك	ثانوي	7	أرمل	سيئة	5744
13	39	أنثى	2	إيجار	فوق الجا	4	متزوج	ممتازة	14302
14	34	ذكر	5	إيجار	جامعي	2	أرمل	ممتازة	12918
15	31	أنثى	4	إيجار	جامعي	0	مطلق	ممتازة	10214
16	47	أنثى	0	إيجار	فوق الجا	5	متزوج	جيدة	11300
17	62	ذكر	4	إيجار	جامعي	8	متزوج	متوسطة	8510
18	43	أنثى	5	إيجار	جامعي	5	مطلق	جيدة	9629
19	65	ذكر	4	ملك	فوق الجا	3	متزوج	ممتازة	12640
20	25	ذكر	5	إيجار	جامعي	4	متزوج	ممتازة	11911
21	25	ذكر	4	إيجار	فوق الجا	6	أعزب	ممتازة	15175
22									

٢- لإيجاد الجدول التكراري لمتغير «الحالة الاقتصادية» قم بتتبع الخطوات التسلسلية من ١ إلى ٥ في برنامج SPSS كما هو موضح في الشكل (٤-٢) أدناه:

شكل رقم (٤-٢)

خطوات إيجاد الجدول التكراري للبيانات غير المبوبة لمتغير «الحالة الاقتصادية» باستخدام SPSS



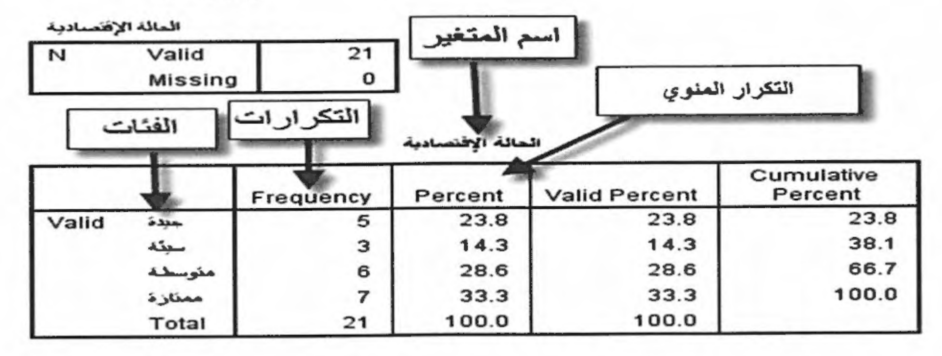
وبعد الضغط على أيقونة «OK» في الخطوة رقم ٥ نحصل على التالي:

شكل رقم (٤-٣)

مخرجات تنفيذ الجدول التكراري غير المبوب لمتغير «الحالة الاقتصادية» باستخدام SPSS

➜ Frequencies

[DataSet1]



Statistics

الحالة الإقتصادية

N	Valid	21
	Missing	0

الحالة الإقتصادية

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	جيدة	5	23.8	23.8	23.8
	سيئة	3	14.3	14.3	38.1
	متوسطة	6	28.6	28.6	66.7
	ممتازة	7	33.3	33.3	100.0
	Total	21	100.0	100.0	

ويمكن إعادة إخراج الجدول في الشـــكل (٤-٣) أعـــلاه عند الحاجة لتضمينه في تقرير البحث أو الدراسة كالتالي:

جدول رقم (٤-٣)

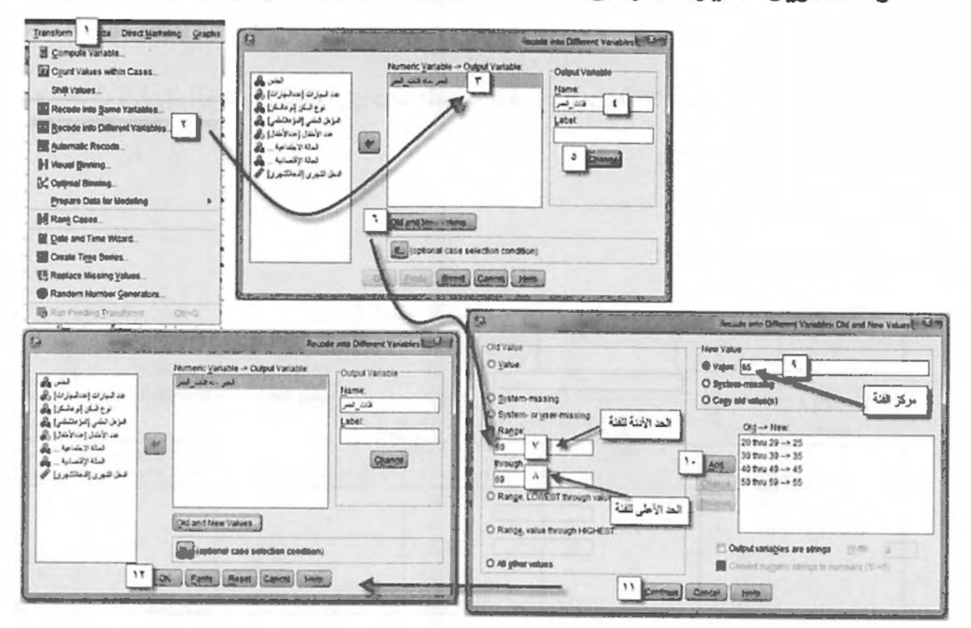
جدول التوزيع التكراري لمتغير الحالة الاقتصادية

الفئة	التكرار	التكرار المئوي
سيئة	٣	١٤,٣
متوسطة	٦	٢٨,٦
جيدة	٥	٢٣,٨
ممتازة	٧	٣٣,٣
المجموع	٢١	١٠٠

٣ - لإيجاد الجدول التكراري لمتغير العمر بعد تحويله إلى فئات من الطول ١٠ سنوات قـــم بتتبع الخطوات التسلســـلية من ١ إلى ١٢ في برنامـــج SPSS كما هو موضح بشكل (٤-٤) أدناه:

شكل رقم (٤-٤)

خطوات تحويل متغير العمر إلى فئات ذات طول ١٠ سنوات لكل فئة باستخدام SPSS



وبعــد الضغط على أيقونة «OK» في الخطوة رقم ١٢ نحصل متغير «فئات العمر» في نافذة محرر البيانات كما هو موضح أدناه:

شكل رقم (٤-٥)

نافذة محرر البيانات وقد احتوت على المتغير الجديد «فئات العمر» الذي تم إنشاؤه باستخدام الخطوات ١-٨

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

متغير "فئات العمر" الذي تم إنشاؤه

25 : الحالة_الصحية

	العمر	الجنس	عدد_السيارات	نوع_المسكن	المؤهل_العلمي	عدد_الأطفال	الحالة_الاجتماعية	الحالة_الصحية	الدخل_الشهري	فئات_العمر
1	37	ذكر	5	ملك	أقل من ث	5	مطلق	متوسطة	9636	30-39
2	54	أنثى	2	ملك	جامعي	1	متزوج	جيدة	12623	50-59
3	65	ذكر	5	ملك	ثانوي	1	مطلق	متوسطة	8996	60-69
4	29	أنثى	5	ملك	ثانوي	6	أعزب	متوسطة	8281	20-29
5	35	ذكر	4	ملك	ثانوي	10	متزوج	سيئة	7090	30-39
6	36	أنثى	3	ملك	جامعي	1	مطلق	متوسطة	6002	30-39
7	43	ذكر	5	إيجار	ثانوي	9	متزوج	سيئة	5208	40-49
8	25	ذكر	5	ملك	ثانوي	2	متزوج	متوسطة	5245	20-29
9	55	ذكر	1	إيجار	جامعي	5	متزوج	جيدة	11406	50-59
10	22	أنثى	4	إيجار	جامعي	4	متزوج	جيدة	10433	20-29
11	42	ذكر	1	إيجار	فوق الجا	8	أرمل	ممتازة	11591	40-49
12	32	أنثى	5	ملك	ثانوي	7	أرمل	سيئة	5744	30-39
13	39	أنثى	2	إيجار	فوق الجا	4	متزوج	ممتازة	14302	30-39
14	34	ذكر	5	إيجار	جامعي	2	أرمل	ممتازة	12918	30-39
15	31	أنثى	4	إيجار	جامعي	0	مطلق	ممتازة	10214	30-39
16	47	أنثى	0	إيجار	فوق الجا	5	متزوج	جيدة	11300	40-49
17	62	ذكر	4	إيجار	جامعي	8	متزوج	متوسطة	8510	60-69
18	43	أنثى	5	إيجار	جامعي	5	مطلق	جيدة	9629	40-49
19	65	ذكر	4	ملك	فوق الجا	3	متزوج	ممتازة	12640	60-69
20	25	ذكر	5	إيجار	جامعي	4	متزوج	ممتازة	11911	20-29
21	25	ذكر	4	إيجار	فوق الجا	6	أعزب	ممتازة	15175	20-29
22										

الآن ســنقوم باتباع نفس الخطوات التسلســلية من ١-٥ كما في الشــكل (٤-٢) للحصــول على الجــدول التكراري لمتغير «فئــات العمر» الموضح في الشــكل (٤-٦) أدناه.

شكل رقم (٤-٦)
مخرجات تنفيذ الجدول التكراري المبوب لمتغير «فئات العمر» باستخدام SPSS

➔ Frequencies

[DataSet1]

Statistics

العمر (Binned)

N	Valid	21
	Missing	0

اسم المتغير

فئات العمر | التكرار | التكرار المئوي

العمر (Binned)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	20-29	5	23.8	23.8	23.8
	30-39	7	33.3	33.3	57.1
	40-49	4	19.0	19.0	76.2
	50-59	2	9.5	9.5	85.7
	60-69	3	14.3	14.3	100.0
	Total	21	100.0	100.0	

ويمكن إعادة إخراج الجدول في الشكل (٤-٦) أعلاه عند الحاجة لتضمينه في تقرير البحث أو الدراسة كالتالي:

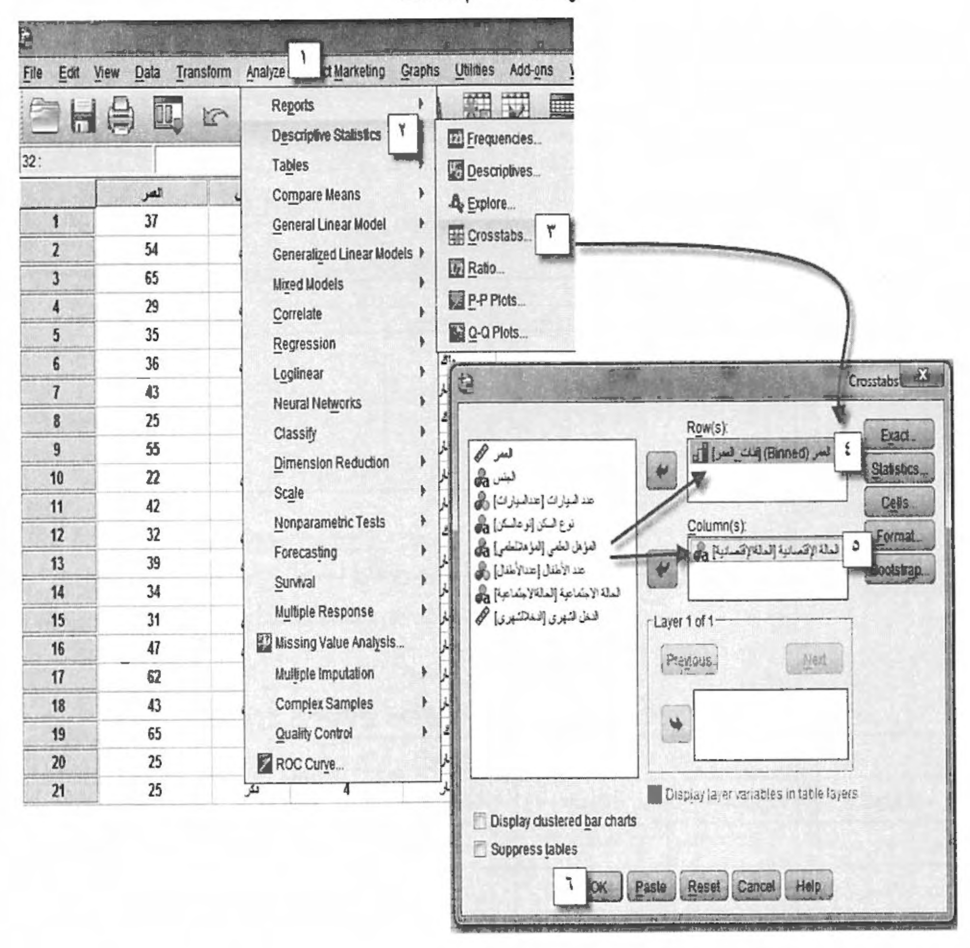
جدول رقم (٤-٤)
جدول التوزيع التكراري لمتغير العمر

الفئة	التكرار	التكرار المئوي
٢٠-٢٩	٥	٢٣,٨
٣٠-٣٩	٧	٣٣,٣
٤٠-٤٩	٤	١٩,٠
٥٠-٥٩	٢	٩,٥
٦٠-٦٩	٣	١٤,٣
المجموع	٢١	١٠٠

٤- لإيجـــاد الجدول التكراري المزدوج لمتغير «الحالـــة الاقتصادية» و«فئات العمر»، قم بتتبع الخطوات التسلســلية من ١ إلى ٦ فـــي برنامج SPSS كما هو موضح أدناه:

شكل رقم (٤-٧)

خطوات إيجاد الجدول التكراري المزدوج لمتغيري «الحالة الاقتصادية» و«فئات العمر» باستخدام SPSS



وبعـــد الضغط علـــى أيقونة «OK» في الخطوة رقم ٦ نحصـــل على جدول التوزيع التكراري المزدوج في شاشة المخرجات كما هو موضح أدناه:

شكل رقم (٤-٨)
مخرجات تنفيذ جدول التوزيع التكراري المزدوج لمتغيري «الحالة الاقتصادية» و«فئات العمر» باستخدام SPSS

Crosstabs

[DataSet1]

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
العمر (Binned) * الحالة الإقتصادية	21	100.0%	0	0.0%	21	100.0%

جدول التوزيع التكراري المزدوج لمتغيري " الحالة الاقتصادية " و "فئات العمر"

العمر (Binned) * الحالة الإقتصادية Crosstabulation

Count

		الحالة الإقتصادية				Total
		جيدة	سيئة	متوسطة	ممتازة	
العمر (Binned)	20-29	1	0	2	2	5
	30-39	0	2	2	3	7
	40-49	2	1	0	1	4
	50-59	2	0	0	0	2
	60-69	0	0	2	1	3
Total		5	3	6	7	21

ويمكن إعادة إخراج الجدول في الشكل (٤-٨) أعلاه عند الحاجة لتضمينه في تقرير البحث أو الدراسة كالتالي:

جدول رقم (٤-٥)
جدول التوزيع التكراري المزدوج لمتغير الحالة الاقتصادية وفئات العمر

فئات العمر	الحالة الاقتصادية				المجموع
	سيئة	متوسطة	جيدة	ممتازة	
٢٠-٢٩	٠	٢	١	٢	٥
٣٠-٣٩	٢	٢	٠	٣	٧
٤٠-٤٩	١	٠	٢	١	٤
٥٠-٥٩	٠	٠	٢	٠	٢
٦٠-٦٩	٠	٢	٠	١	٣
المجموع	٣	٦	٥	٧	٢١

الرسومات الإحصائية Statistical Graphs:

إن الرســومات الإحصائيــة أو الأشــكال البيانيــة هي أحد طــرق وصف وتحليل مجموعــات البيانات أو المتغيــرات بطريقة بيانية. فمن خلالها يســتطيع الباحث أو القــارئ التعرف على شــكل وطبيعة توزيع البيانات، المقارنة بــين مجموعات البيانات والظواهــر، العلاقة بين المتغيــرات، الاتجاه العام للظواهر خــلال الزمن- وكما قيل: «فــإن الصــورة تعبر عن ألف كلمة». وهي أســهل في توصيل المعلومــات إلى القارئ والمستفيد وتساعده في الحصول على بعض القيم والأرقام بشكل سريع ودون الحاجة للحسابات أو الصيغ الرياضية.

ومن الرسومات الإحصائية الأكثر استخداماً في البحوث والتقارير ما يلي:

١- الأعمدة البيانية Bar Charts: وتستخدم لتمثيل البيانات بيانياً في حالة المتغيرات النوعية ذات المستوى الاسمي أو الرتبي. والأعمدة البينانية نوعان هما:

- الأعمدة البيانية البســيطة Simple Bar Chart وذلك لتمثيل متغير واحد فقط بيانياً.
- الأعمدة البيانية المركبة وتأخذ صورتين مختلفتين هما الأعمدة البيانية المجمعة Clustered Bar Charts والأعمــدة البيانية المكدســة Stacked Bar Charts وتستخدم في حالة تمثيل متغيرين أو أكثر بيانياً.

وفي الواقع العملي يفضل اســتخدام الأعمــدة البيانية في حالة المقارنات بين المتغيرات كالمقارنة بين أعــداد الموظفين الذكور والإناث في مجموعة من الأجهزة الحكومية.

٢- الدائرة البيانية Pie Chart: وتســتخدم لتمثيل البيانات بيانياً في حالة المتغيرات النوعية ذات المستوى الاسمي أو الرتبي، كما هو الحال بالنسبة للأعمدة البيانية. ولكــن فــي الواقع العملي يفضــل اســتخدامها للتعرف على توزيــع متغير واحد فقــط حيث إنها تعطي صورة أوضح من الأعمدة البيانيــة من ناحية المقارنة بين فئات المتغير. وكما هو واضح من تسميتها فهي عبارة عن دائرة مقسمة إلى أجزاء أو قطاعات بأحجام متناسبة مع تكرار أو نسبة كل جزء أو قطاع في المجموعة.

٣- الخط البياني أو السلســلة الزمنية Line Chart or Sequence Chart: ويستخدم هذا النوع من الرســومات الإحصائية لتمثيل بيانات الظاهرة محل الدراسة خلال الزمــن. وهــذا يفيد فــي التعرف على النمــط والاتجاه العــام (صاعد، هابط،

متذبذب) للظاهرة خلال الزمن. والخط البياني قد يكون بسيط Simple Line لتمثيل ظاهرة أو متغير واحد فقط أو متعدد Multiple Lines لتمثيل ظاهرتين أو متغيرين أو أكثر في آن واحد خلال الزمن. ومثال على ذلك أعداد العمالة الوافدة والمغادرة للمملكة العربية السعودية للأعوام من ٢٠١٠م إلى ٢٠١٣م.

٤- المدرج التكراري Histogram: يتم استخدام المدرج التكراري لتمثيل المتغيرات الكمية بيانياً بعد تحويلها إلى فئات ويتكون المدرج التكراري من محورين متعامدين ومستطيلات متلاصقة. المحور الأفقي يمثل فئات المتغيرين والمحور الرأسي يمثل التكرارات، وكل مستطيل يرتكز على فئة من فئات المتغير حيث يمثل ارتفاعه تكرار الفئة. ويستخدم المدرج التكراري للتعرف بشكل رئيسي على شكل توزيع الظاهرة أو المتغير محل الدراسة من حيث التماثل حول نقطة مركزية ودرجة الالتواء Skewness ودرجة التفلطح أو التحدب Kurtosis، وهل تحتوي البيانات قيماً شاذة أم لا.

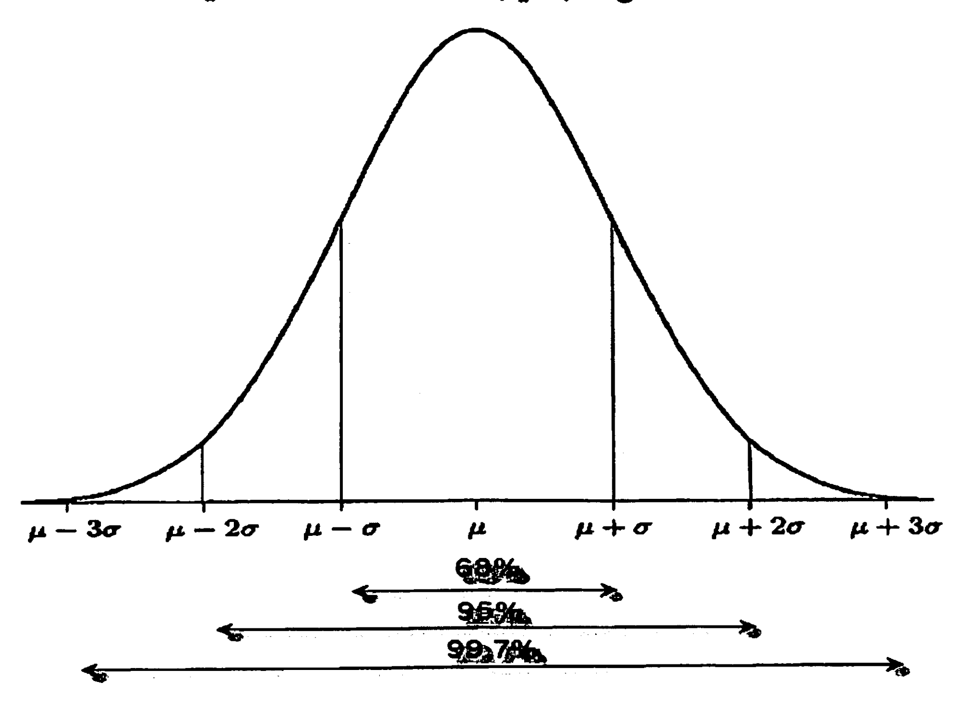
٥- المنحنى التكراري Frequency Curve: ويستخدم لتمثيل البيانات الكمية ذات المقياس الفتري أو النسبي بيانياً. ويستخدم لنفس الأغراض للمدرج التكراري. ومن منحنيات التوزيع التكراري أو الاحتمالي الأكثر شهرة واستخداماً منحنى التوزيع الطبيعي ومنحنى t. ونظراً لأهمية التوزيع الطبيعي Normal Distribution ولأنه سيرد ذكره في مواضع عديدة من الكتاب، فإنه من الضروري التعرف على مفهومه وأهميته. إن منحنى التوزيع الطبيعي هو من أشهر التوزيعات الاحتمالية وأكثرها استخداماً. وهو عبارة عن منحنى متماثل يشبه الشكل الجرسي لذا يطلق عليه أحياناً منحنى الشكل الجرسي Bell-Shaped Curve، ويتم تحديده بمعلمتين فقط هما المتوسط الحسابي (μ) والذي يعتبر أحد وأشهر مقاييس النزعة المركزية، وكذلك الانحراف المعياري (σ) والذي يعتبر أيضاً أحد وأشهر مقاييس التشتت أو التباين للبيانات - انظر الشكل (٤-٩). فإذا كان هناك متغير ما يتبع التوزيع الطبيعي كأوزان مجموعة من الأشخاص مثلاً، فإن هذا يعني أن معظم تلك الأوزان تتمركز حول المتوسط الحسابي لتلك الأوزان في حين أن القليل جداً من تلك الأوزان تبتعد عن المتوسط سواء بالزيادة او النقصان.

ومن الحقائق الخاصة والهامة بالتوزيع الطبيعي، هو أنه إذا كان توزيع مجموعة من البيانات لمتغير ما طبيعي أو شبه طبيعي (القانون التجريبي Empirical Rule) فإن:

- (٦٨٪) من تلك البيانات يتمركز حول متوسطها الحسابي وتنتشر في فترة Interval حول المتوسط بمقدار واحد انحراف معياري عن المتوسط.
- (٩٥٪) من تلك البيانات يتمركز حول متوسطها الحسابي وتنتشر في فترة بمقدار اثنين انحراف معياري عن المتوسط.
- (٩٩,٧٪) من تلك البيانات يتمركز حول متوسطها الحسابي وتنتشر في فترة بمقدار ثلاثة انحراف معياري عن المتوسط.

شكل رقم (٤-٩)

منحنى التوزيع الطبيعي بمتوسط μ وانحراف معياري σ



وتكمن أهمية التوزيع الطبيعي في أن معظم الظواهر الاجتماعية، الطبيعية، الفيزيائية، الحيوية، الاقتصادية، الإدارية وغيرها تتبع التوزيع الطبيعي تقريباً، وكمثال على ذلك الذكاء، الطول، الوزن والدخل...إلخ . كما أن له استخدامات عدة في مجال الأساليب الإحصائية لمراقبة وضبط الجودة. بالإضافة إلى ذلك فإن التوزيع الطبيعي يعتبر هاماً وركيزة أساسية لمعظم أساليب الاستدلال

الإحصائي. علاوة على ذلك فإن هناك نظرية في الإحصاء (نظرية النهاية المركزية Central Limit Theorem) تنص على أن توزيع المتوسط الحسابي لعينات من نفس الحجم المسحوبة من أي مجتمع بغض النظر عن شكل توزيعه التكراري أو الاحتمالي يتبع تقريباً التوزيع الطبيعي عندما يكون حجم العينة كبيراً. مع التأكيد على أنه إذا كانت العينات مسحوبة من مجتمع طبيعي فإن توزيع المتوسط الحسابي للعينة يتبع تماماً التوزيع الطبيعي لجميع أحجام العينات.

٦- الشكل الانتشاري Scatter plot: يستخدم لتمثيل العلاقة بين متغيرين كميين بيانياً. فمن خلاله يمكن التعرف على طبيعة واتجاه العلاقة بين متغيرين كميين، بمعنى هل العلاقة بين المتغيرين خطية أم غير خطية أم أنه لا يوجد علاقة بينهما، وفي حال وجود علاقة خطية مثلاً بين المتغيرين فهل هي طردية أم عكسية. في الرسم الانتشاري ينبغي أن تكون بيانات المتغيرين أزواج مرتبة Paired Data. والشكل الانتشاري يتكون من محورين متعامدين ونقاط منتشرة في المستوى المحصور بين المحورين تمثل بيانات المتغيرين محل الدراسة. ولقد جرت العادة على أن يمثل المحور الأفقي المتغير المستقل والمحور الرأسي يمثل المتغير التابع إن أمكننا تمييزهما كمتغير مستقل وتابع.

كيفية إنشاء الرسومات الإحصائية باستخدام برنامج SPSS من خلال تطبيق عملي تفاعلي:

سيتم توضيح كيفية إنشاء الرسومات الإحصائية باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٤-٢):

بالجوع إلى التطبيق العملي التفاعلي (٤-١)، سيتم تنفيذ ما يلي باستخدام برنامج SPSS:

١- مقارنة الذكور والإناث «متغير الجنس» تبعاً لمتغير «نوع السكن» ملك أم إيجار بيانياً.

٢- إنشاء الدائرة البيانية لمتغير الحالة الاقتصادية.

٣- إنشاء المدرج التكراري لمتغير «فئات العمر»، موضحاً عليه المنحنى التكراري.

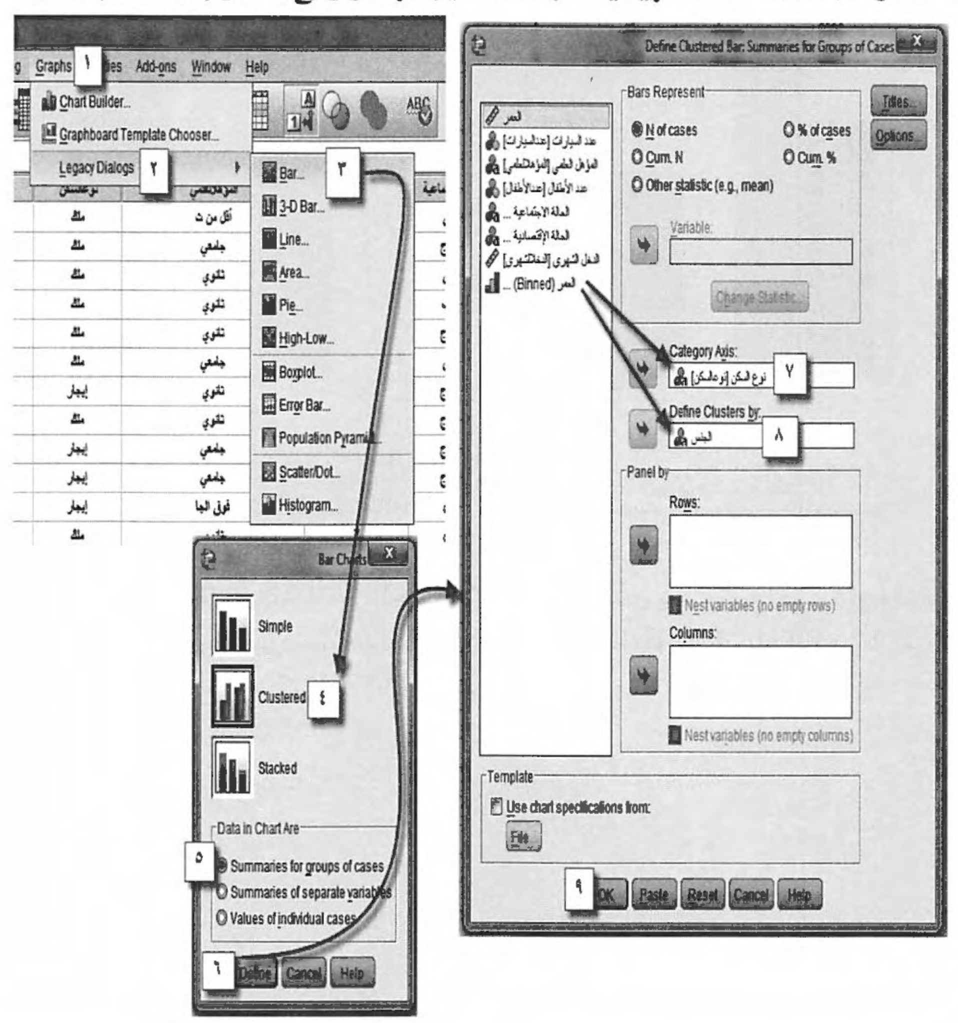
الحل:

١- لمقارنة الذكور والإناث «متغير الجنس» تبعاً لمتغير «نوع السكن» ملك أم إيجار بيانياً، سيتم استخدام الأعمدة البيانية المجمعة Clustered Bar Chart لكون

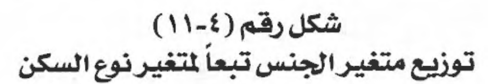
المتغيرين نوعيين من المستوى الاســـمي. ولتنفيذ ذلك باستخدام SPSS قم بتتبع الخطوات التسلسلية من ١ إلى ٩ كما هو موضح بالشكل (٤-١٠) أدناه:

شكل رقم (٤-١٠)

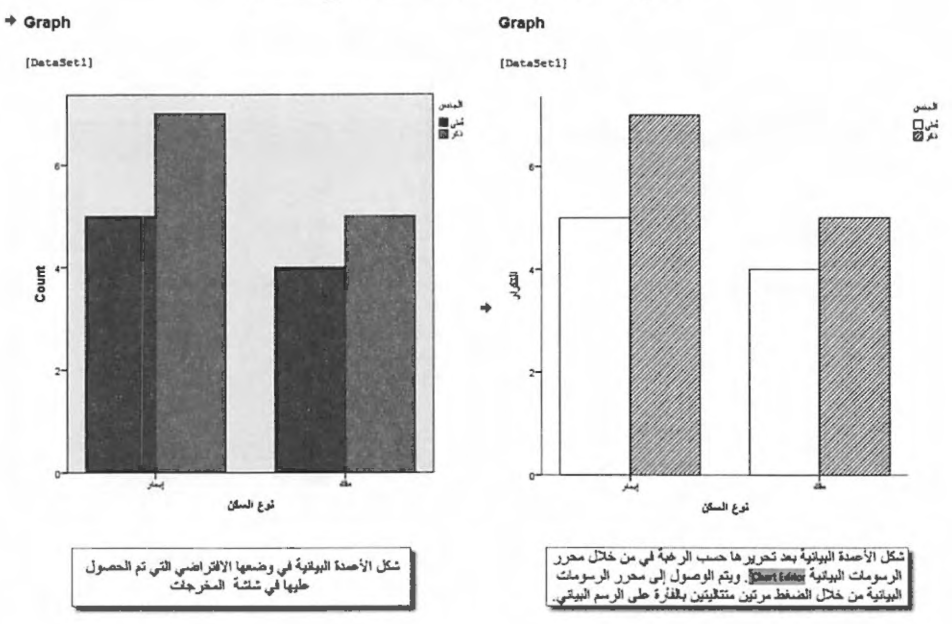
خطوات إنشاء الأعمدة البيانية المجمعة لمتغير الجنس ونوع السكن باستخدام SPSS



وبعد التنفيذ من خلال الضغط على أيقونة «OK» في الخطوة رقم ٩، يتم الحصول على التالي في شاشة المخرجات:



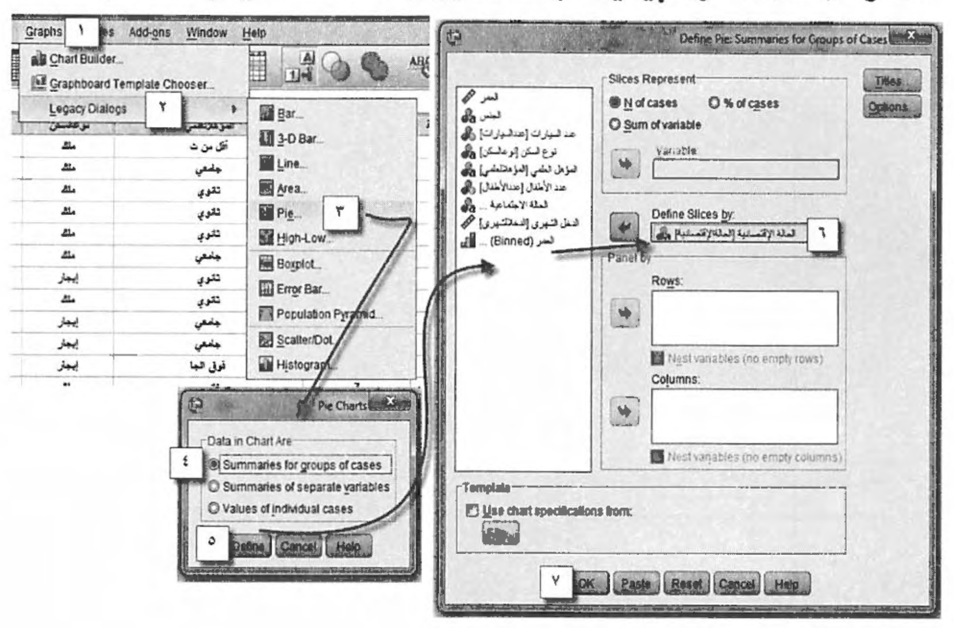
شكل رقم (٤-١١)
توزيع متغير الجنس تبعاً لمتغير نوع السكن



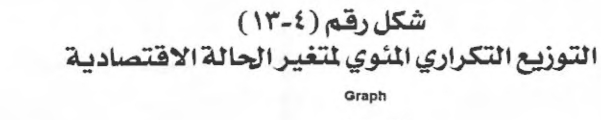
٢- لإنشاء الدائرة البيانية لمتغير الحالة الاقتصادية والذي يعدُّ متغيراً نوعياً أو وصفياً من المســتوى الرتبي باستخدام SPSS، قم بتتبع الخطوات التسلسلية من ١ إلى ٧ كما هو موضح بالشكل (٤-١٢) أدناه:

شكل رقم (٤-١٢)

خطوات إنشاء الدائرة البيانية المجمعة لمتغير الحالة الاقتصادية باستخدام SPSS

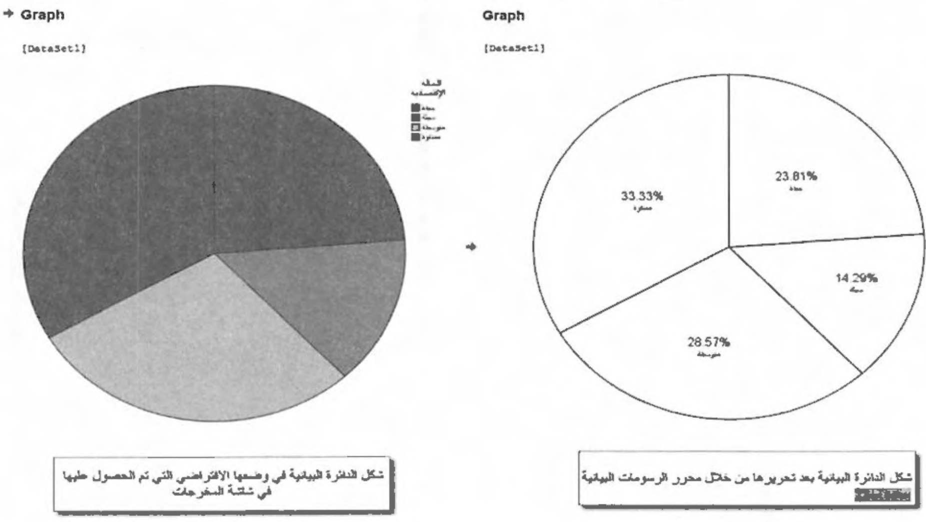


وبعد التنفيذ من خلال الضغط على أيقونة «OK» في الخطوة رقم ٧، يتم الحصول على التالي في شاشة المخرجات:



شكل رقم (٤-١٣)

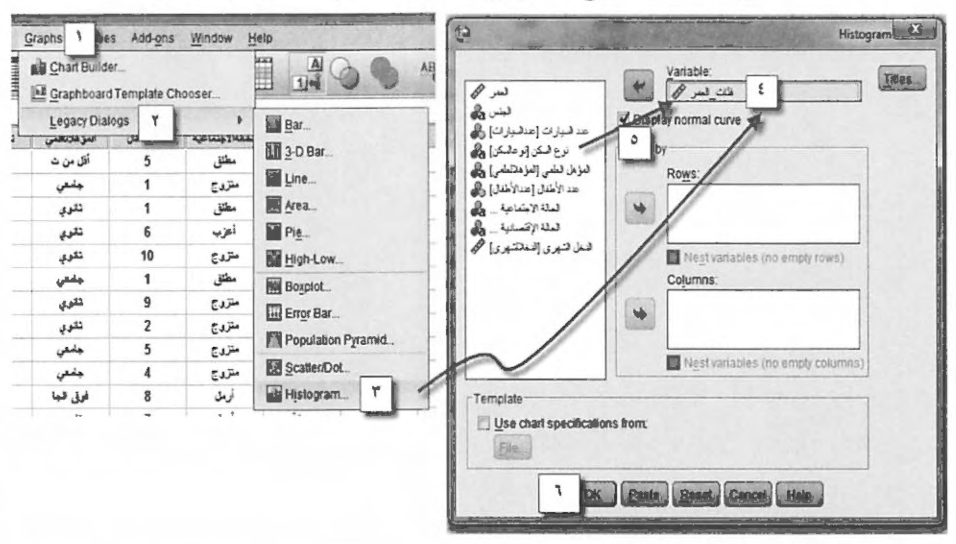
التوزيع التكراري المئوي لمتغير الحالة الاقتصادية



٣- لإنشاء المدرج التكراري لمتغير فئات العمر موضحاً عليه المنحنى التكراري باستخدام SPSS، قم بتتبع الخطوات التسلسلية من ١ إلى ٦ كما هو موضح بالشكل (٤-١٤) أدناه:

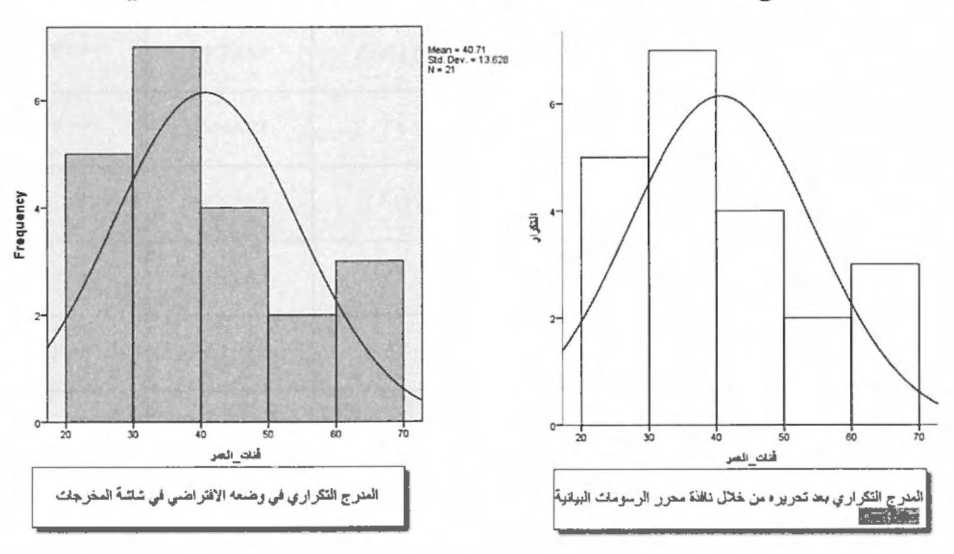
شكل رقم (٤-١٤)

خطوات إنشاء المدرج التكراري لمتغير العمر باستخدام SPSS



وبعد التنفيذ من خلال الضغط على أيقونة «OK» في الخطوة رقم ٦، يتم الحصول على التالي في شاشة المخرجات:

شكل رقم (٤-١٥)
التوزيع التكراري لمتغير فئات العمر موضحاً عليه المنحنى التكراري



تطبيق عملي تفاعلي (٤-٣):

الجدول (٤-٦) يمثل صادرات المملكة العربية السعودية من وإلى الخارج بملايين الريالات للأعوام من ١٩٨٨م إلى ٢٠٠٧م، وذلك حسب ما ورد في موقع مصلحة الإحصاءات العامة والمعلومات.

جدول رقم (٤-٦)
صادرات وواردات المملكة العربية السعودية (بملايين الريالات) للأعوام من ١٩٨٨ -٢٠٠٧م

العام	الصادرات	الواردات	العام	الصادرات	الواردات
١٩٨٨	٩١٢٨٨	٨١٦٠٧	١٩٩٨	١٤٥٣٨٨	١١٢٣٩٧
١٩٨٩	١٠٦٢٩٤	٧٩٢٧٨	١٩٩٩	١٩٠٠٨٤	١٠٤٩٨٠
١٩٩٠	١٦٦٣٩٩	٩٠٢٨٢	٢٠٠٠	٢٩٠٥٥٣	١١٣٢٤٠

تابع - جدول رقم (٤-٦).

العام	الصادرات	الواردات	العام	الصادرات	الواردات
١٩٩١	١٧٨٦٣٦	١٠٨٩٣٤	٢٠٠١	٢٥٤٨٩٨	١١٦٩٣١
١٩٩٢	١٨٨٣٢٥	١٢٤٦٠٦	٢٠٠٢	٢٧١٧٤١	١٢١٠٨٨
١٩٩٣	١٥٨٧٧٠	١٠٥٦١٦	٢٠٠٣	٣٤٩٦٦٤	١٥٦٣٩١
١٩٩٤	١٥٩٥٩٠	٨٧١٩٢	٢٠٠٤	٤٧٢٤٩١	١٧٧٦٥٩
١٩٩٥	١٨٧٤٠٣	١٠٥١٨٧	٢٠٠٥	٦٧٧١٤٤	٢٢٢٩٨٥
١٩٩٦	٢٢٧٤٢٨	١٠٣٩٠٨	٢٠٠٦	٧٩١٣٣٩	٢٦١٤٠٢
١٩٩٧	٢٢٧٤٤٣	١٠٧٦٤٣	٢٠٠٧	٨٧٤٤٠٣	٣٣٨٠٨٨

المطلوب:

استخدام برنامج SPSS لتنفيذ ما يلي:

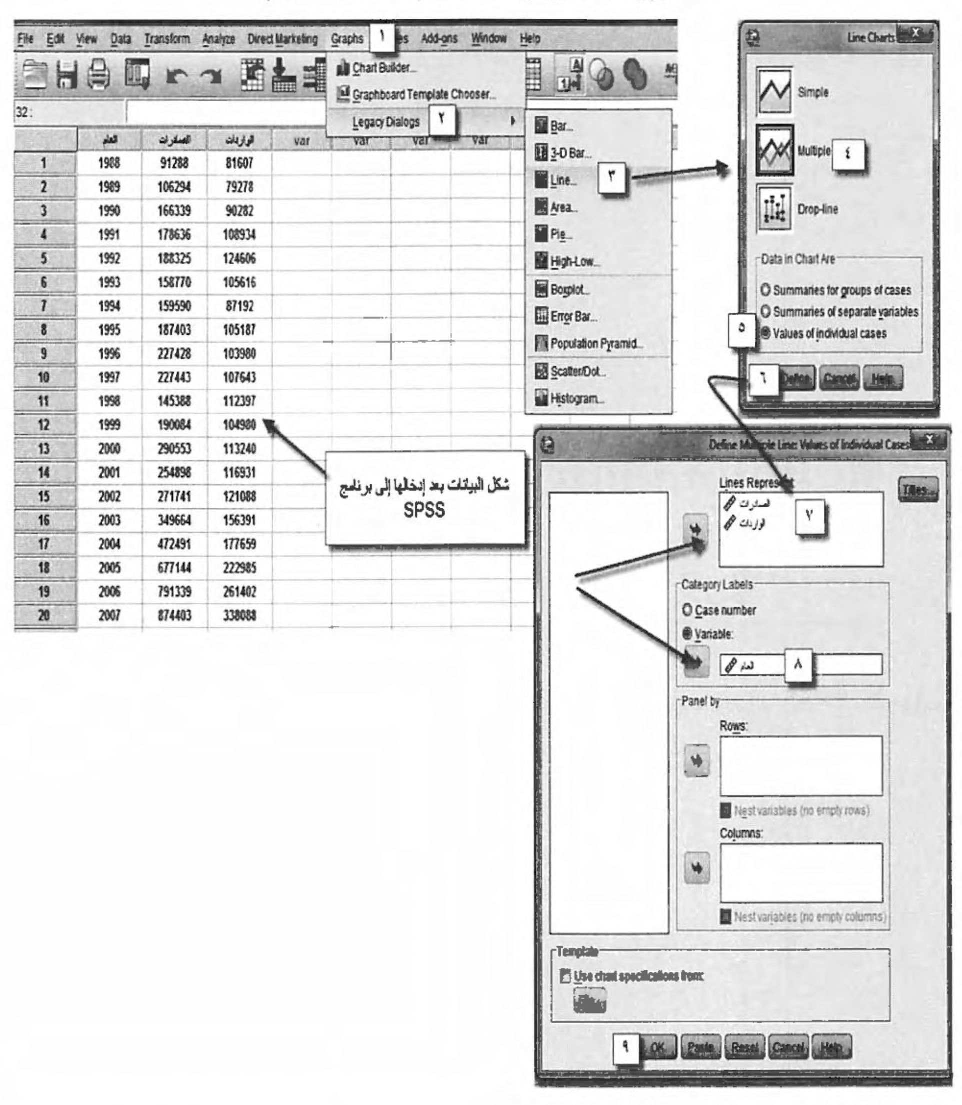
١- إنشاء الخط البياني أو السلسلة البيانية المتعددة لمتغيري الصادرات والواردات.

٢- إنشاء الشكل الانتشاري للصادرات والواردات.

الحل:

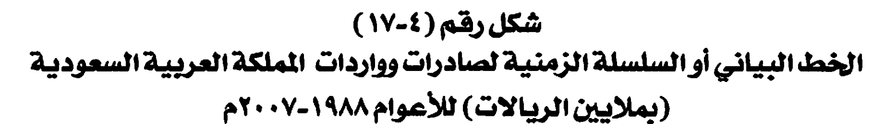
١- لإنشاء الخط البياني المتعدد لمتغيري الصادرات والواردات خلال الزمن باستخدام SPSS، قم بتتبع الخطوات التسلسلية من ١ إلى ٩ كما هو موضح في الشكل (٤-١٦) أدناه:

شكل رقم (٤-١٦)
خطوات إنشاء الخط البياني أو السلسلة البيانية لمتغيري صادرات وواردات المملكة العربية السعودية للأعوام ١٩٨٨-٢٠٠٧م



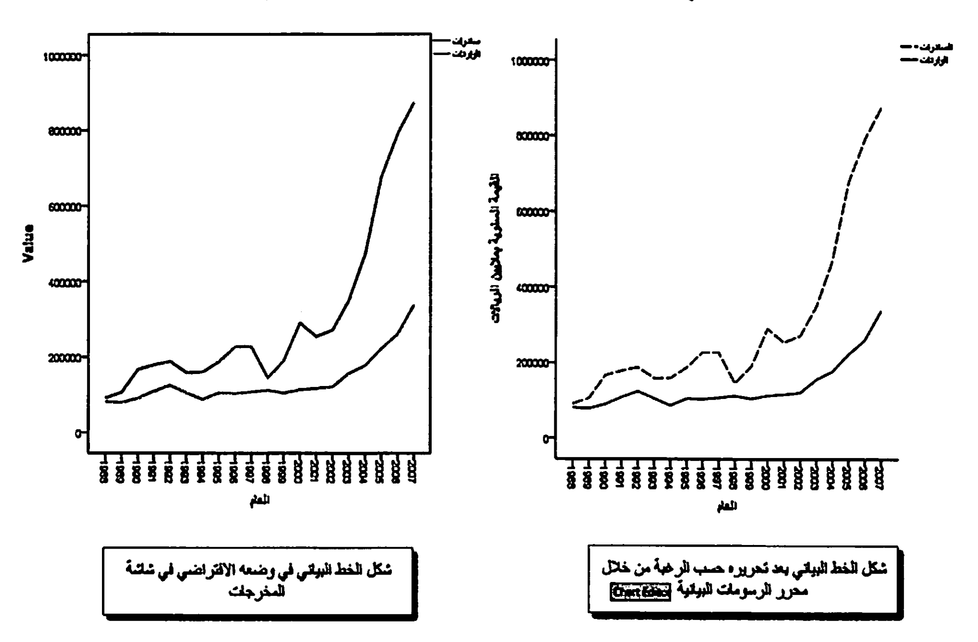
	العام	الصادرات	الواردات
1	1988	91288	81607
2	1989	106294	79278
3	1990	166339	90282
4	1991	178636	108934
5	1992	188325	124606
6	1993	158770	105616
7	1994	159590	87192
8	1995	187403	105187
9	1996	227428	103980
10	1997	227443	107643
11	1998	145388	112397
12	1999	190084	104980
13	2000	290553	113240
14	2001	254898	116931
15	2002	271741	121088
16	2003	349664	156391
17	2004	472491	177659
18	2005	677144	222985
19	2006	791339	261402
20	2007	874403	338088

وبعد التنفيذ من خلال الضغط على أيقونة «OK» في الخطوة رقم ٩، يتم الحصول على التالي في شاشة المخرجات:



شكل رقم (٤-١٧)

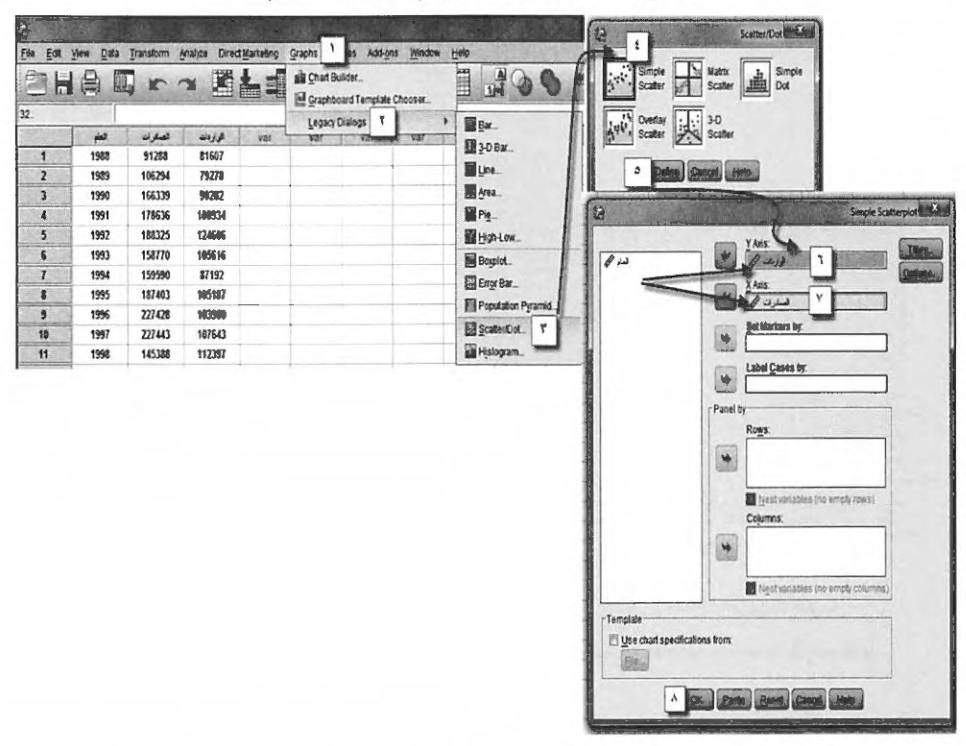
الخط البياني أو السلسلة الزمنية لصادرات وواردات المملكة العربية السعودية (بملايين الريالات) للأعوام ١٩٨٨-٢٠٠٧م



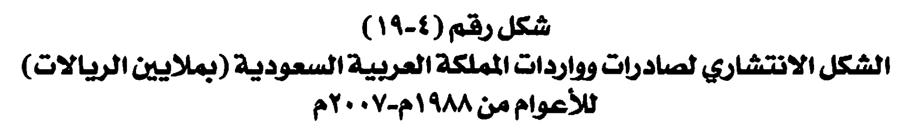
٢- لإنشاء الشكل الإنتشاري لمتغيري الصادرات والواردات باستخدام SPSS، قم بتتبع الخطوات التسلسلية من ١ إلى ٨ كما هو موضح في الشكل (٤-١٨) أدناه:

شكل رقم (٤-١٨)

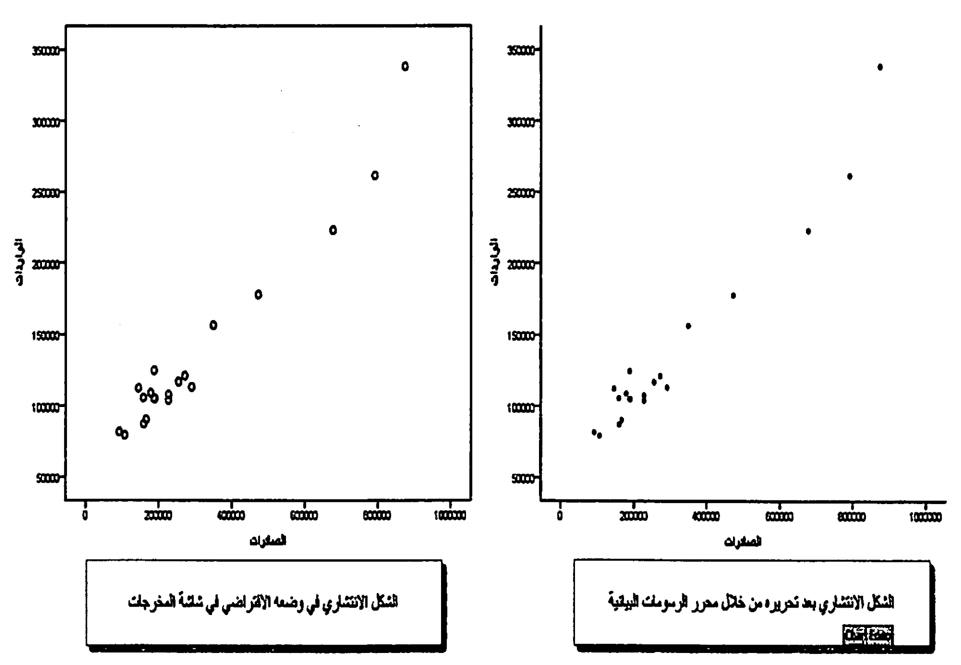
خطوات إنشاء الخط البياني أو السلسلة البيانية لمتغيري صادرات وواردات المملكة العربية السعودية للأعوام ١٩٨٨-٢٠٠٧م



وبعد التنفيذ من خلال الضغط على أيقونة «OK» في الخطوة رقم ٨، يتم الحصول على التالي في شاشة المخرجات:



شكل رقم (٤-١٩)
الشكل الانتشاري لصادرات وواردات المملكة العربية السعودية (بملايين الريالات) للأعوام من ١٩٨٨م-٢٠٠٧م



مقاييس النزعة المركزية Measures of Central Tendency:

مما لاشـــك فيه أن الجداول التكرارية والرسومات البيانية تعتبر هامة في التعرف على طبيعة وشكل توزيع المتغيرات والظواهر محل الدراسة ووصفها وإبراز خصائصها والعلاقـــات بينها. ولكـــن تظل عملية وصف المتغيرات والظواهر من خلال الأشـــكال البيانية محل تباين في وجهات النظر بين المستفيدين والقراء فهي لا تستند إلى مقاييس كميـــة أو قوانين رياضية يســـتطيع اثنان من خلالها الوصول إلى نفس الاســـتنتاجات بسهولة. فمثلاً، لا نستطيع بشكل قطعي تحديد القيمة التي تتمركز أو تتجمع حولها البيانات أو تحديد مقدار التشـــتت أو التباين في البيانات من خلال الأشكال البيانية، وإنما يمكن إعطاء أرقام كمية تقريبية تختلف من فرد لآخر حســـب وجهة نظره. لذا نشـــأت الحاجة لبناء مقاييس كمية تســـاعد في تزويد الباحث والمستفيد في الوصول إلى مؤشرات كمية لا تختلف قيمها من شخص إلى آخر.

ومن المؤشـــرات أو المقاييـــس الكمية الهامة جـداً في وصف البيانـــات، مقاييس النزعة المركزية وتســـمى في كثير من الأحيان - وكما هـــو متعارف عليه بين كثير من المستخدمين - بالمتوسطات Averages. وتعرف مقاييس النزعة المركزية أو المتوسطات علـــى أنها تلك المقاييـــس التي تعطي قيماً تتمركـــز وتتجمع حولها معظـــم البيانات. ومقاييس النزعة المركزية كثيرة منها المتوســـط الحســـابي، الوسيط، المنوال، المتوسط الهندسي، المتوسط التوافقي وسنكتفي هنا بالحديث عن أكثر تلك المقاييس استخداماً في البحوث والدراسات وهي المتوسط الحسابي، الوسيط، والمنوال. ويعرض الجدول (٤-٧) ملخصـــاً لتلك المقاييس من حيث التعريف والخصائص والاســـتخدام. وتجدر الإشارة إلى أنه يمكن حساب مقاييس النزعة المركزية للبيانات المبوبة وغير المبوبة.

جدول رقم (٤-٧)

مقارنة بين مقاييس النزعة المركزية: المتوسط، الوسيط، المنوال

المقياس	تعريفه	نوع المتغير ومستوى قياسه	مزاياه	عيوبه
المتوسط Mean	حاصل مجموع القيم على عددها.	متغير كمي من المستوى الفئوي أو النسبي	- يعتمد على كل القيم في حسابه. - سهل الفهم والحساب جبرياً. - يدخل في حساب كثير الأساليب الإحصائية الأخرى. - يفضل في حالة التوزيعات المتماثلة (غير الملتوية) أو القريبة من التماثل.	- تتأثر قيمته بالقيم الشاذة أو المتطرفة. - لا يمكن إيجاده للبيانات النوعية.
الوسيط Median	القيمة التي تقع في منتصف البيانات بعد ترتبها تصاعدياً أو تنازلياً	- متغير كمي من المستوى الفئوي أو النسبي - متغيرنوعي من المستوى الرتبي.	- لا يتأثر بالقيم الشاذة أو المتطرفة. - يمكن إيجاده في حالة البيانات النوعية ذات القياس الرتبي. - يفضل في حالة التوزيعات الملتوية.	- لا يعتمد في حسابه على كل البيانات محل الدراسة. - لا يدخل في حساب كثير من الأساليب الإحصائية الأخرى.

تابع - جدول رقم (٤-٧).

المقياس	تعريفه	نوع المتغير ومستوى قياسه	مزاياه	عيوبه
المنوال Mode	القيمة الأكثر تكراراً في مجموعة البيانات	المتغيرات الكمية والنوعية	- لا يتأثر بالقيم الشاذة أو المتطرفة. - يمكن إيجاده للبيانات الكمية والنوعية.	- قد يوجد أكثر من منوال أو قد لا يوجد منوال لمجموعة من البيانات. - لا يدخل في حساب الأساليب الإحصائية الأخرى.

مقاييس التشتت (التباين) Dispersion (Variation) Measures:

إن مقاييس النزعة المركزية تزودنا بالقيمة التي تتمركز حولها البيانات للمتغير أو الظاهرة محل الدراسة، ولكن لا نستطيع من خلال هذه المقاييس التعرف على مدى تشتت البيانات وتناثرها حول القيمة المتوسطة والتي تعتبر مهمة جداً في دراسة الظواهر والمتغيرات. إننا نستطيع التعرف على تباين المتغيرات وتشتتها من خلال ما يعرف بمقاييس التشتت أو التباين، والتي تعرف على أنها تلك المقاييس التي توضح كمياً مدى تشتت البيانات بعضها عن بعض أو مدى تباعد وتقارب البيانات عن نقطة التمركز (المتوسط). ومفهوم التباين هام جداً في الإحصاء وعليه تقوم أساليب التحليل الاحصائي. وهناك مقاييس عديدة لتباين أو تشتت البيانات منها المدى، الانحراف المتوسط، التباين والانحراف المعياري، المدى الربيعي، معامل الاختلاف، الاختلاف النسبي. والجدول (٤-٨) يقدم أشهر تلك المقاييس وخصائصها.

جدول رقم (٤-٨)

مقارنة بين مقاييس التشتت: المدى، الانحراف المعياري (والتباين)، المدى الربيعي، معامل الاختلاف، ونسبة الاختلاف

المقياس	تعريفه	نوع المتغير ومستوى قياسه	مزاياه	عيوبه
المدى Range	حاصل الفرق بين أعلى قيمة وأصغر قيمة في البيانات.	متغير كمي من المستوى الفئوي أو النسبي.	- سهل الفهم والحساب. - من تطبيقاته العملية: الجودة والأحوال الجوية حيث إنه يركز فقط الحد الأعلى والأدنى للظاهرة.	- يتأثر قيمته بالقيم الشاذة أو المتطرفة. - يعتمد على قيمتين فقط في حسابه. - لا يمكن إيجاده للبيانات النوعية.
الانحراف المعياري (والتباين) Standard Deviation (Variance)	الجذر التربيعي لمتوسط مربعات انحرافات القيم عن متوسطها الحسابي.	متغير كمي من المستوى الفئوي أو النسبي.	- يعتمد على كل القيم في حسابه. - يدخل في حساب كثير الأساليب الإحصائية الأخرى. - يفضل في حالة التوزيعات المتماثلة (غير الملتوية) أو القريبة من التماثل.	- تتأثر قيمته بالقيم الشاذة أو المتطرفة. - لا يمكن إيجاده للبيانات النوعية. - أصعب في الحساب الجبري والفهم من المدى.
المدى الربيعي Inter-Quartile Range	حاصل الفرق بين الربيع الأول Q1 والربيع الثالث Q3 للبيانات بعد ترتيبها تصاعدياً أو تنازلياً.	المتغيرات الكمية من المستوى الفئوي أو النسبي.	- لا يتأثر بالقيم الشاذة أو المتطرفة. - يمكن إيجاده في حالة البيانات النوعية ذات القياس الرتبي. - يفضل في حالة التوزيعات الملتوية.	- لا يعتمد في حسابه على كل البيانات محل الدراسة. - لا يدخل في حساب كثير من الأساليب الإحصائية الأخرى.

تابع - جدول رقم (٤-٨).

المقياس	تعريفه	نوع المتغير ومستوى قياسه	مزاياه	عيوبه
معامل الاختلاف Coefficient of Variation	(١) - هو حاصل قسمة الانحراف المعياري على المتوسط الحسابي مضروباً في ١٠٠. كلما اقتربت قيمته من الصفر دل على التجانس الكبير في البيانات والعكس.	المتغيرات الكمية من المستوى الفئوي أو النسبي.	- يعتمد على كل القيم في حسابه. - يفضل في حالة التوزيعات المتماثلة (غير الملتوية) أو القريبة من التماثل. - يستخدم لمقارنة نسبة التباين بين مجموعتين من البيانات مختلفة في الوحدات أو مختلفة في متوسطاتها. - يبين لنا ما إذا كانت نسبة التشتت في البيانات صغيرة نسبياً أم كبيرة نسبياً. - يستخدم في تقييم نماذج الانحدار الخطي.	
	(٢) - هو حاصل قسمة المدى الربيعي على الوسيط مضروبا في ١٠٠. كلما اقتربت قيمته من الصفر دل على التجانس الكبير في البيانات والعكس.	- ا لمتغيرا ت الكمية من المستوى الفئوي أو النسبي - ا لمتغيرا ت النوعية من المستوى الرتبي.	- يستخدم لمقارنة نسبة التباين بين مجموعتين من البيانات مختلفة في الوحدات أو مختلفة في متوسطاتها. - يبين لنا ما إذا كانت نسبة التشتت في البيانات صغيرة نسبياً أم كبيرة نسبياً. - يفضل استخدامه في حالة التوزيعات شديدة الالتواء.	- لا يعتمد على كل القيم في حسابه.

تابع - جدول رقم (٤-٨).

المقياس	تعريفه	نوع المتغير ومستوى قياسه	مزاياه	عيوبه
نسبة الاختلاف Variation Ratio	حاصل قسمة «عدد التكرارات خارج الفئة المنوالية» على «مجموع التكرارات الكلية». وهو موجب وقيمته أقل من الواحد الصحيح. فإذا كانت قيمة مقياس نسبة الاختلاف قريبة من الصفر فإنه يقال بأن البيانات متجانسة أي تقع كلها تقريباً أو معظمها في الفئة المنوالية والعكس صحيح. أي أنه كلما ابتعدت قيمته عن الصفر فإن البيانات متشتتة وتتوزع على فئات المتغير.	المتغيرات النوعية.	- إيجاد قيمة التشتت كمياً في حالة البيانات النوعية.	

مقاييس شكل التوزيع Shape of Distribution Measures:

إن مؤشري النزعة المركزية والتشتت للمتغير أو الظاهرة محل الدراسة ليسا كافيين لاكتمال الصورة حول سلوك توزيع المتغير بل نحتاج إلى معرفة خاصية ثالثة عن توزيع ذلك المتغير ألا وهي شكل التوزيع والمتمثل في خاصيتين هما: الالتواء Skewness والتفرطح Kurtosis. إن معرفة شكل توزيع البيانات يعتبر هاماً جداً في التحليل الإحصائي بل يعتبر من أولى الخطوات التي ينبغي أن يقوم بها الباحث قبل البدء في عملية التحليل الإحصائي حيث إنه يساعد الباحث على التعرف وبشكل أوضح وأسرع على شكل توزيع البيانات ومدى خلوها من القيم الشاذة والمتطرفة ومن ثم اختيار الأسلوب الإحصائي المناسب في التحليل. إن وصف شكل التوزيع البياني لمتغير الدراسة بالعين المجردة فقط يعتبر طريقة جدلية غير موضوعية تختلف من شخص لآخر حسب وجهة نظره وخبرته. لذا نشأت الحاجة إلى تصميم مقاييس إحصائية موضوعية تعطي قيماً كمية للبيانات محل الدراسة لا يختلف عليها اثنان.

الالتواء Skewness:

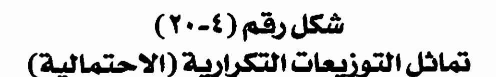
يقال لتوزيع بيانات متغير ما أنه متماثلاً Symmetric إذا كان متوسط ذلك التوزيع يقسم المنحنى أو المدرج التكراري إلى قسمين متماثلين أو متطابقين. ومن أشهر التوزيعات الاحتمالية المتماثلة التوزيع الطبيعي أو الاعتدالي وتوزيع t. ويمكننا من خلال المدرج التكراري أوالمنحنى التكراري التعرف بيانياً على ما إذا كان شكل توزيع المتغير أو الظاهرة محل الدراسة متماثلاً أو غير متماثل (ملتو)، ولكن لا نستطيع من الشكل البياني تحديد كمية أو مقدار الالتواء في البيانات. إن تحديد مقدار أو كمية التواء المتغيرات تعتبر غاية في الأهمية في التحليل الإحصائي حيث أنه يعطينا دلالة على توزيع البيانات حول متوسطها، بمعنى هل تتوزع البيانات بالتساوي على يمين ويسار المتوسط أم أنها تتركز في إحدى الجهتين أكثر من غيرها. كما أن درجة شدة الالتواء تحدد نوع الأسلوب الإحصائي في التحليل، فمثلاً في حالة التوزيعات التكرارية (الاحتمالية) المتماثلة أو شبه المتماثلة يمكننا استخدام المتوسط الحسابي لوصف مركز البيانات، أما إذا كانت البيانات غير متماثلة أو شديدة الالتواء فإنه يفضل استخدام الوسيط.

معامل الالتواء:

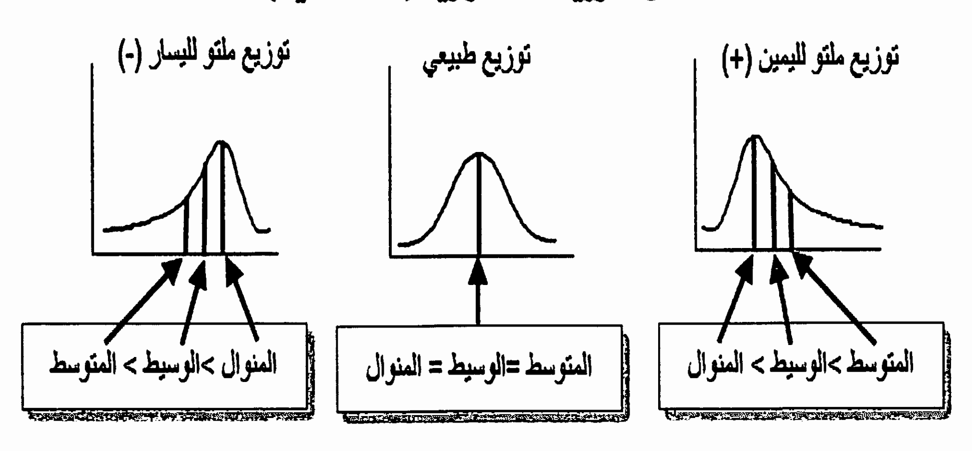
هو مقياس أو مؤشر إحصائي لتقدير مقدار الالتواء أو الانحراف عن التماثل في المتغير أو الظاهرة محل الدراسة كمياً. ويعرف رياضياً لمجموعة من بيانات عينة $x_1, x_2, \ldots, x_n$ حجمها n ومتوسطها الحسابي $\bar{x}$ وانحرافها المعياري s كالتالي :(Snedecor & Cochran, 1989, 79)

$$Skewness = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{(n - 1)s^3}$$

إن معامل الالتواء للتوزيع الطبيعي يساوي الصفر، وكذلك الحال بالنسبة لأي توزيع متماثل. أما إذا كانت قيمة معامل الالتواء أكبر من الصفر (أي موجبة) فإن ذلك يدل على أن هناك التواء (عدم تماثل) في توزيع البيانات اتجاه اليمين (الاتجاه الموجب) وعندها يقال إن التوزيع ملتو لليمين أو ذو التواء موجب فهذا يعني أن معظم البيانات أقل من المتوسط، أما إذا كانت قيمة معامل الالتواء أقل من الصفر (أي سالبة) والتي عندها يقال أن التوزيع ملتو لليسار أو ذو التواء سالب فإن ذلك يدل على أن هناك التواء (عدم تماثل) في توزيع البيانات اتجاه اليسار (الاتجاه السالب)، وهذا يعني أن معظم البيانات أعلى من المتوسط. وتزداد شدة الالتواء كلما ابتعدت قيمة معامل الالتواء عن الصفر. والشكل (٤-٢٠) يوضح التوزيع الطبيعي المتماثل وتوزيعات أخرى ملتوية.



شكل رقم (٤-٢٠)
تماثل التوزيعات التكرارية (الاحتمالية)



وكما يتضح من الشكل (٤-٢٠) أعلاه فإن العلاقة بين مقاييس النزعة المركزية (المتوسط، الوسيط، المنوال) تعتمد على تماثل واتجاه الالتواء في توزيع البيانات. فمثلاً إذا كان المتوسط والوسيط لمجموعة من البيانات متساوياً أو قريبة من بعضها بشكل كبير جداً فإن توزيع البيانات يكون متماثلاً أو شبه متماثل، وهنا يمكن للباحث الاختيار بين المتوسط أو الوسيط، لكن يفضل استخدام المتوسط نظراً لمألوفيته لدى عامة الناس ولكونه يدخل في مكونات الأساليب الإحصائية الأخرى.

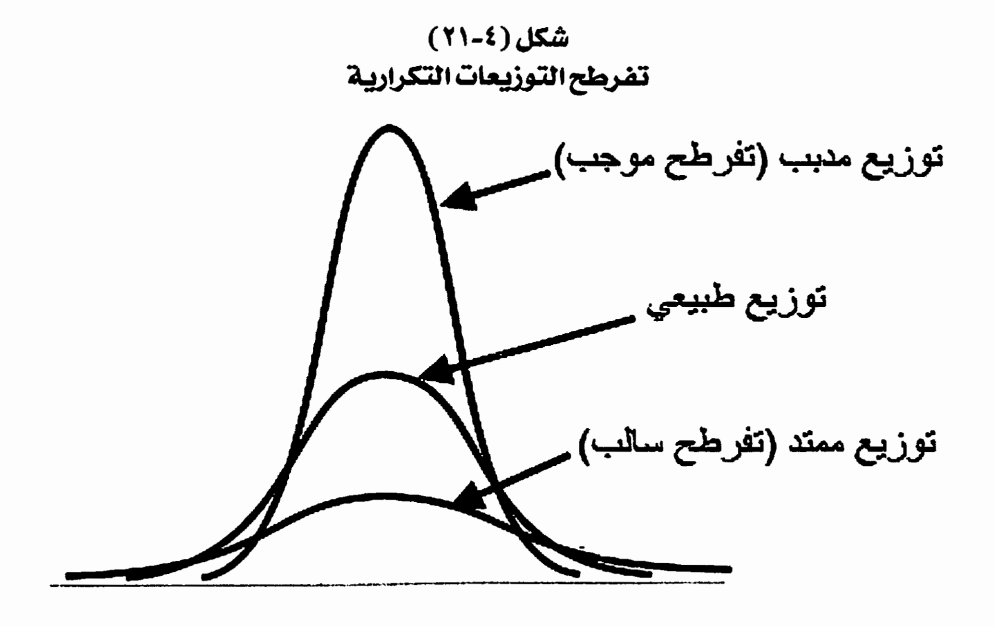
معامل التفرطح (التمدد) Kurtosis:

هو مقياس أو مؤشر إحصائي لتقدير مقدار التمدد أو الانحراف في المتغير أو الظاهرة محل الدراسة كمياً. ويعرف رياضياً لمجموعة من بيانات عينة $x_1, x_2, \ldots, x_n$ حجمها n ومتوسطها الحسابي $\bar{x}$ وانحرافها المعياري s كالتالي (Snedecor & Cochran, 1989, 79):

$$Kurtosis = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{(n - 1)s^4} - 3$$

إن معامل التفرطح للتوزيع الطبيعي يساوي الصفر. فإذا إذا كانت قيمة معامل التفرطح أكبر من الصفر، فإن هذا يعني أن قمة المنحنى أو المدرج التكراري تكون أعلى من التوزويع الطبيعي، وعندها يقال بأن التوزيع مدبب peaked أو ذو تفرطح موجب، ومن ثم نستنتج أن البيانات تتمركز وتتجمع حول مركز (متوسط) البيانات بشكل أكبر من التوزيع الطبيعي. أما إذا كانت قيمة معامل التفرطح أقل من الصفر فإن هذا يعني أن قمة المنحنى التكراري للبيانات أقل من التوزيع الطبيعي، وعندئذ يقال بأن التوزيع ممتد flat أو ذو تفرطح سالب، ومن ثم نستنتج أن البيانات تتمركز أو تتجمع حول متوسطها بشكل أقل من التوزيع الطبيعي وتميل للانتشار تجاه أطراف التوزيع. والشكل (٤-٢١) التالي يوضح التوزيع الطبيعي والتفرطح الموجب والسالب.

شكل (٤-٢١)
تفرطح التوزيعات التكرارية



ونلاحظ مما سبق ما يلي:

- إذا كان التوزيع طبيعياً فإن معامل الالتواء له يساوي الصفر والعكس ليس بالضرورة صحيحاً على إطلاقه، وذلك لأن معامل الالتواء يساوي صفراً للتوزيعات المتماثلة. فمثلاً توزيع t هو توزيع متماثل معامل الالتواء له يساوي الصفر، ولكنه لا يتبع التوزيع الطبيعي.
- إذا كان التوزيع طبيعياً فإن معامل التفرطح له يساوي الصفر والعكس ليس بالضرورة صحيحاً على إطلاقه.

إضاءة إحصائية حول الخصائص الوصفية اللازمة لوصف المتغير أو الظاهرة محل الدراسة:

في الحياة العملية، عندما يتم جمع بيانات حول ظاهرة معينة فإن الاستفادة منها في وضعها الخام يكون محدود الفائدة أو شبه منعدم، ولاسيما عندما يكون حجم البيانات كبيراً ويصعب علينا التعرف على سلوك توزيع تلك البيانات والوصول إلى معلومات واستنتاجات قيمة ومفيدة. وللوصول إلى صورة واضحة المعالم حول توزيع بيانات متغير الدراسة فإنه لابد من معرفة وتقدير الخصائص الثلاثة التالية للبيانات، وهي:

- مركــز البيانات. حيث ينبغي تقدير القيمة المتوســطة التي تتمركز حولها البيانات ويتم ذلك من خلال استخدام أحد مقاييس النزعة المركزية.
- تشتت أو تباين البيانات. إن معرفة متوسط البيانات غير كاف لأخذ تصور كاف عن توزيع البيانات، فقد يكون لدينا مجموعتان من البيانات لهما نفس المتوسط ولكنهما تختلفان في تشتت البيانات حول متوسطها. لذلك فإنه ينبغي استخدام مؤشر آخر يبين لنا مقدار التشــتت في تلك البيانات كمياً، وهذا يتم من خلال اســتخدام أحد مقاييس التشتت.
- شــكل التوزيع. إن معرفة متوسط البيانات وتشتتها غير كاف لاكتمال الصورة حول توزيــع بيانــات المتغير. فمن المهم جــداً معرفة ما إذا كان توزيــع البيانات متماثلاً طبيعيــاً أو متماثلاً غير طبيعي أو ملتوياً، وكذلك ما إذا كان التوزيع مدبباً أو ممتداً جهــة الأطراف، بالإضافة إلى ما إذا كان للتوزيع منــوال (قمة) واحد فقط أم أكثر من منوال.

إضاءة إحصائية حول طرق دراسة تماثل البيانات:

هناك عدة طرق للحكم على تماثل بيانات المتغير الكمي محل الدراسة منها:

١- الرســم البياني ومن أشهرها المدرج والمنحنى التكراري. يكون التوزيع متماثلاً إذا أمكن تقســيم الشكل إلى جزأين متساويين بحيث ينطبق أحدهما على الآخر حول خط رأسي يمر بمركز البيانات.

٢- مقارنة مقاييس النزعة المركزية. فإذا كان المتوسط الحسابي والوسيط متساويين أو متقاربــين إلى حد كبير جداً فإن التوزيــع متماثل. عدا ذلك فهو غير متماثل، ومــن ثم يكون ملتويــاً جهة اليمين إذا كان المتوســط أكبر من الوســيط والعكس صحيح.

٣- معامل الالتواء.

كيفيــة إيجاد مقاييس النزعة المركزية والتشــتت وشــكل التوزيع باستخدام برنامج SPSS من خلال تطبيق عملي تفاعلي:

ســيتم توضيح كيفية إيجاد المتوســطات ومقاييس التشتت وشــكل توزيع البيانات باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٤-٤):

بالرجــوع إلــى التطبيق العملي التفاعلي (٤-١)، ســيتم تنفيذ ما يلي باســتخدام برنامج SPSS.

١- إيجاد قيمة المتوسط الحسابي، الوسيط، والمنوال لمتغير الدخل الشهري.

٢- إيجاد قيمة المدى، والانحراف المعياري، المدى الربيعي لمتغير الدخل الشهري.

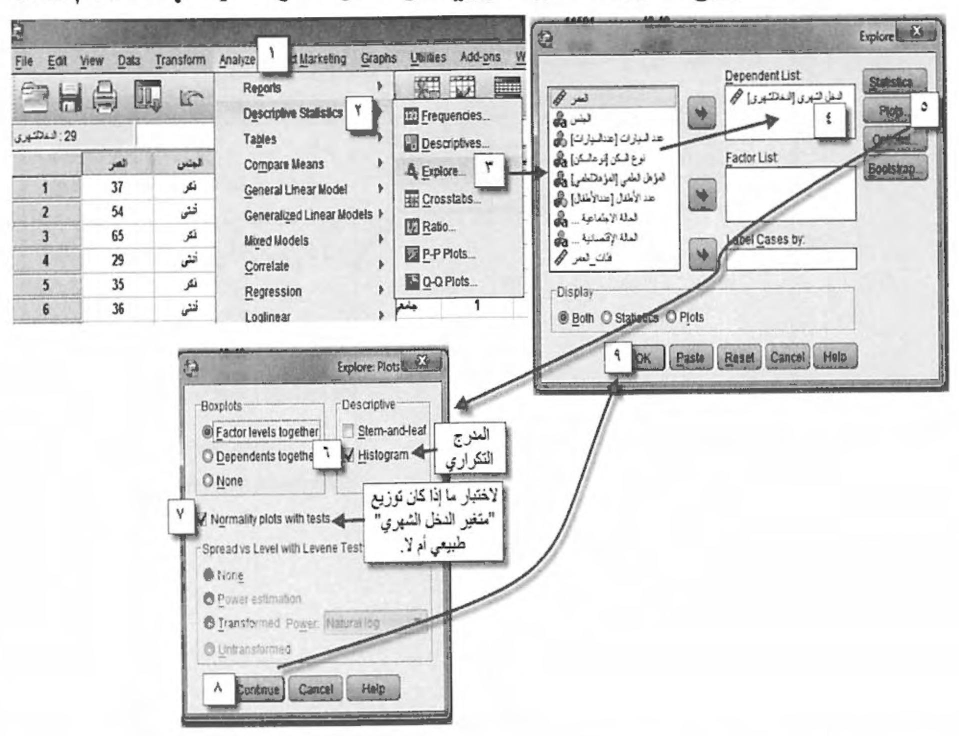
٣- إيجاد قيمة معاملي الالتواء والتفرطح لمتغير الدخل الشهري، وتفسيرهما.

الحل:

١- للحصول على المطلوب في الفقرات ١، ٢، ٣ ســيتم استخدام الأمر Explore والذي يعنى «يستكشــف» في برنامج SPSS، ولتنفيذ ذلك قم بتتبع الخطوات التسلسلية من ١ إلى ٩ كما هو موضح بالشكل (٤-٢٢) أدناه:

شكل (٤-٢٢)

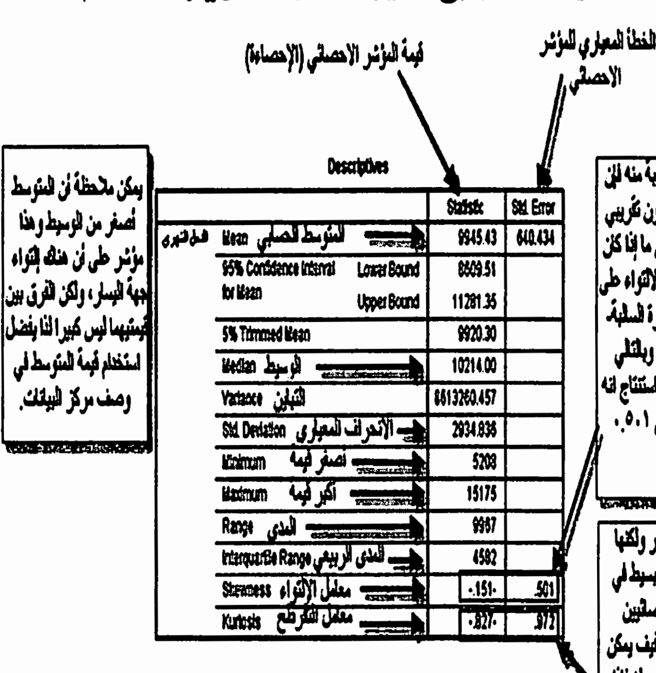
خطوات استكشاف توزيع متغير «الدخل الشهري» من خلال الأمر Explore باستخدام SPSS



وبعد التنفيذ من خلال الضغط على أيقونة «Ok» في الخطوة رقم ٩، يتم الحصول على التالي في شاشة المخرجات:

شكل (٤-٢٣)

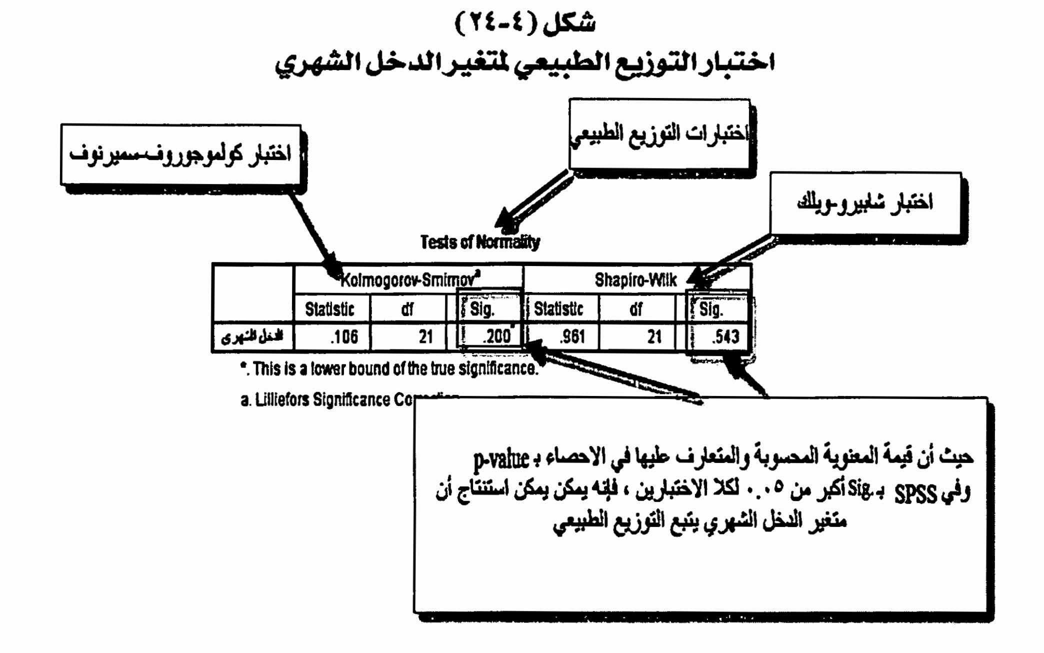
قيم النزعة المركزية والتشتت وشكل التوزيع لمتغير الدخل الشهري باستخدام SPSS



Descriptives

			Statistic	Std. Error
الدخل الشهري	Mean المتوسط الحسابي		9945.43	640.434
	95% Confidence Interval for Mean	Lower Bound	8609.51	
		Upper Bound	11281.35	
	5% Trimmed Mean		9920.30	
	Median الوسيط		10214.00	
	Variance التباين		8613260.457	
	Std. Deviation الإنحراف المعياري		2934.836	
	Minimum أصغر قيمة		5208	
	Maximum أكبر قيمة		15175	
	Range المدى		9967	
	Interquartile Range المدى الربيعي		4582	
	Skewness معامل الإلتواء		-.151-	.501
	Kurtosis معامل التفرطح		-.827-	.972

شكل (٤-٢٤)
اختبار التوزيع الطبيعي لمتغير الدخل الشهري



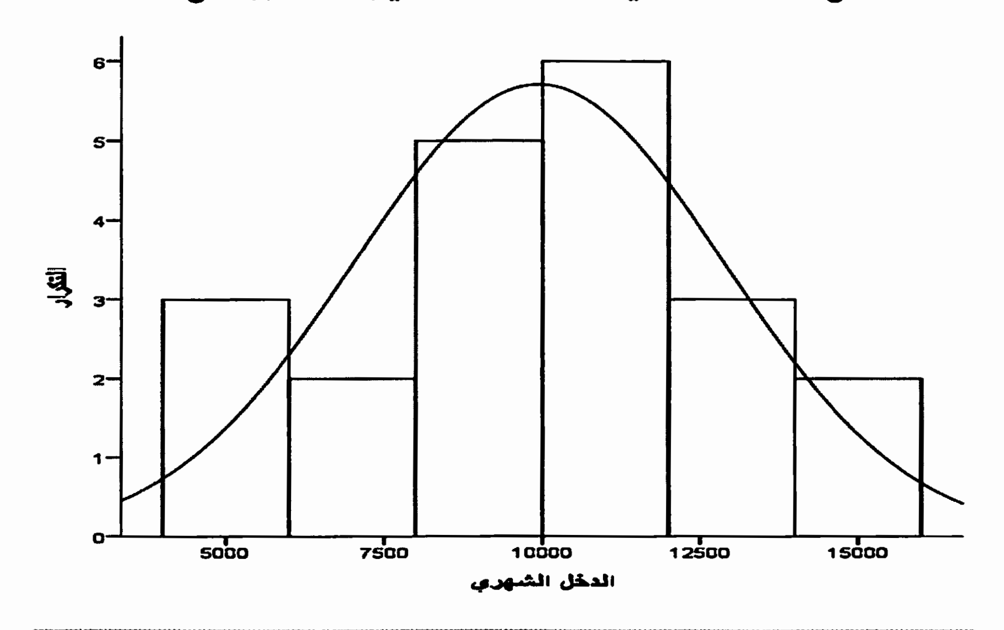
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
الدخل الشهري	.106	21	.200*	.961	21	.543

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

شكل (٤-٢٥)
المدرج والمنحنى التكراري لمتغير الدخل الشهري باستخدام برنامج SPSS



الاستنتاج:

كما هو موضح بالأشــكال (٤-٢٣)، (٤-٢٤)، و(٤-٢٥) يمكننا اســتنتاج أن متغير الدخل الشــهري بالريال الســعودي يتبع التوزيع الطبيعي بمتوسط (٩٩٤٥,٤٣) ريال وانحراف معياري (٢٩٣٤,٨٤) ريالاً.

الفصل الخامس

اختبارات الفرضيات الإحصائية Statistical Hypotheses Tests

مقدمة:

في الفصل الأول من الكتاب تم الإشارة إلى أن أسئلة الدراسة وفرضياتها تنبثق من أهدافها. حيث يجب أن تكون تلك الأسئلة والفرضيات قابلة للدراسة وأن تترجم أهداف الدراسة إجرائياً. وللإجابة عن تساؤلات الدراسة واختبار فرضياتها البحثية بطريقة علمية مبنية على البيانات يتم استخدام الأساليب والاختبارات الإحصائية المناسبة حسب الدراسة. وبناء عليه يتم إعادة صياغة فرضيات الدراسة بأسلوب إحصائي يمكننا من اختبار تلك الفرضيات، وهو ما يعرف بالفرضيات الإحصائية.

الفرضيات الإحصائية Statistical Hypotheses:

الفرض الإحصائي هو ادعاء أو تصريح أو تخمين (قد يكون صحيحاً أو خاطئاً) حول معلمة أو أكثر غير معروفة من معالم مجتمع أو مجتمعات الدراسة يتم التحقق منه بطريقة إحصائية. بمعنى آخر هو تخمين أو إجابة أو تفسير مؤقت يعطي تصوراً عن معلمة من معالم المجتمع أو طبيعة العلاقة بين متغيرين أو أكثر بشكل قابل للاختبار. ونظراً لأهمية صياغة الفرضيات، فإن هناك عدة خصائص ينبغي أن تتصف بها الفرضية المصاغة. وقد تطرق عبيدات، عدس وعبدالحق (٢٠٠٢) وعليان وغنيم (٢٠٠٨) إلى بعض من تلك الخصائص منها:

- معقولية الفرضية، وذلك بأن تتسق الفرضية مع الحقائق المعروفة والنظريات العلمية أو نتائج الدراسات والبحوث السابقة.
- أن يتم صياغة الفرضية قبل مرحلة جمع البيانات، وذلك لأن الفرضية بعد صياغتها هي التي ستوجه وتوفر جهود الباحث من حيث عملية جمع البيانات المتعلقة بالفرضيات وتحديد أساليب وإجراءات التحليل الإحصائي لاختبار الفرضيات وتفسير النتائج.
- أن تقدم الفرضية إجابة أو تفسيراً محتمل لسؤال البحث مما يمكن الباحث من التزام فرضيته بأسئلة الدراسة.
- أن تصاغ الفرضية بعبارات واضحة ومحددة بشكل يستلزم تحديد المفاهيم والمتغيرات التي تشتمل عليها الفرضية بدقة.

- أن توضح الفرضية بشكل دقيق العلاقة المتوقعة بين متغيرات الدراسة المستقلة والتابعة.
- أن تكون الفرضية قابلة للاختبار الإحصائي، وذلك بأن تصاغ الفرضية بطريقة قابلة للقياس ويمكن التحقق منها إجرائياً.
- بساطة الفرضية، وذلك:
- أن تكون للفرضية إجابة «صحيحة» واحدة فقط.
- أن تكون الفرضية في نطاق إمكانات الباحث من حيث الوقت، والمال، والجهد المبذول لاختبار الفرضية.
- عدم دمج أكثر من فرضية منفصلة في فرضية واحدة مركبة.

تصاغ فرضية البحث إحصائياً في صورة فرضيتين متنافيتين تعرفان بفرضيتي العدم والبديلة. ويفضل عادة أن تصاغ فرضية الدراسة في صورة الفرضية البديلة.

فرضية العدم (الفرضية الصفرية أو فرضية الاختبار) Null Hypothesis:

ويرمز لها بالرمز H_0 وتصاغ عادة في صورة عدم وجود اختلاف أو فروق ذات دلالة إحصائية أو عدم وجود علاقة ذات دلالة إحصائية بين المتغيرات ويتم الاحتفاظ بفرضية العدم ولا يتم رفضها إلا في حال توفر الدليل من بيانات العينة، ويتم ذلك عن طريق أحد أساليب التحليل الإحصائي الاستدلالي (الاختبارات الإحصائية أو ما يسمى باختبارات الفرضيات).

الفرضية البديلة (فرضية البحث أو الباحث) Alternative (Research) Hypothesis:

ويرمز لها بالرمز H_1 أو H_a وتصاغ عادة بشكل مخالف لما هو محدد في فرضية العدم، أي يفترض وجود اختلاف جوهري أو علاقة ذات دلالة إحصائية بين المتغيرات. ويتم قبوله في حال رفض فرضية العدم. وتجدر الإشارة إلى أن رفض فرضية العدم لا يعني أنها خاطئة، كذلك فإن عدم رفضها لا يعني بالضرورة أنها صحيحة، وإنما الرفض وعدم الرفض يعود إلى مدى دعم بيانات العينة لفرضية العدم. والفرضية البديلة يمكن لها أن تكون في اتجاه واحد (موجهة) أو في اتجاهين (غير موجهة) حسب ما يراه الباحث، وهذا بدوره يؤدي إلى كون الاختبار الإحصائي في اتجاه واحد One-tailed Test أو في اتجاهين Two-tailed Test، (Moore & Notz, 2006, 452-453).

إن فرضيتـــي العدم والبديلة متلازمتان ومكملتـــان كل منهما للأخرى، ولكن تعتبر فرضية البدل هي الأســـاس في الدراســـات البحثية، فهي التي يتم صياغتها في نص البحوث كونها تحوي اهتمام الباحث من حيث وجود الفروق أو الارتباط بين المتغيرات. أما فرضية العدم فليس هناك حاجة لإدراجها ضمن سياق البحث، ولكن تكمن أهميتها في إجراء اختبار الفرضيات إحصائياً.

تطبيق عملي تفاعلي (٥-١):

في دراســـة ميدانية لقياس «الرضا العام للمستفيدين من خدمات الأحوال المدنية بمدينة الرياض» كان من ضمن أسئلة الدراسة ما يلي:

هـــل يوجد اختلاف ذو دلالة إحصائية في مســـتوى الرضا العام للمســـتفيدين من خدمات الأحوال المدنية بالرياض حسب فئات العمر؟

المطلوب:

صياغة فرضيتي العدم والبدل لسؤال الدراسة السابق.

الإجابة:

H_0: لا يوجد اختلاف ذو دلالة إحصائية (أو معنوية) في متوسط مستوى الرضا العام للمستفيدين من خدمات الأحوال المدنية في الرياض باختلاف فئات العمر.

H_1: يوجد اختلاف ذو دلالة إحصائية (أو معنوية) في متوســـط مستوى الرضا العام للمستفيدين من خدمات الأحوال المدنية في الرياض باختلاف فئات العمر.

ملاحظة:

الفرضيـــة البديلة الســـابقة تعتبر فرضية غير موجهة، وذلـــك لأننا لم نفترض أن متوســـط مســـتوى الرضا يزيد أو ينقص لفئـــة عمرية مقارنة بفئـــة أو فئات عمرية أخرى.

تمرين تطبيقي (٥-١):

لنفرض أن لدينا التساؤل البحثي التالي:

هل تزيد إنتاجية الموظف كلما زاد مستوى رضاه عن بيئة عمله؟

المطلوب:

١- صياغة فرضيتي العدم والبدل لسؤال البحث السابق.

٢- نوع الفرضية البديلة من حيث الاتجاه.

٣- ناقش الفرضية البديلة من حيث مدى تحقيقها لخصائص أو صفات الفرضية الجيدة.

الاختبار الإحصائي Statistical Test:

عبـــارة عن مجموعة من الخطوات والإجراءات يتم مـــن خلالها الوصول إلى قرار حول فرضية العدم، وذلك بناء على قياسات كمية يتم حسابها من بيانات العينة وأخرى محددة من قبل الباحث.

إحصائية الاختبار Test Statistic:

عبارة عن قياس كمي يتم حســـابه من بيانات عينة الدراســـة ليتم اســـتخدامه في اختبار فرضية العدم، وتتغير قيمة إحصائية الاختبار من عينة لأخرى. أي أن إحصائية الاختبار متغير عشوائي له توزيع احتمالي.

أنواع الأخطاء في اختبارات الفرضيات الإحصائية:

عند إجراء الاختبار الإحصائي لاتخاذ قرار حول فرضية العدم بقبولها أو عدم قبولها، فإنه ينتج لدينا أربع حالات ممكنة متعلقة بصحة القرار تعتمد على العاملين التاليين:

١- كون فرضية العدم H_0 صحيحة أو غير صحيحة في الواقع.

٢- قبول أو عدم قبول فرضية العدم H_0 بناء على الاختبار الإحصائي.

وبشـــكل مفصل فإن الحالات الممكنة حول القرار الـــذي يتوصل إليه الباحث حول اختبار فرضية العدم هي:

- الحالة الأول: عدم قبول (رفض) فرضية العدم H_0 عندما تكون صحيحة في الواقع، وهذا قرار خاطئ ويسمى في الإحصاء خطأ من النوع الأول Type I Error.
- الحالة الثانية الثاني: عدم رفض (قبول) فرضية العدم H_0 وهي في الواقع صحيحة، وهذا قرار صحيح.
- الحالة الثالث: عدم رفض (قبول) فرضية العدم H_0 وهي في الواقع غير صحيحة، وهذا قرار خاطئ ويسمى في الإحصاء خطأ من النوع الثاني Type II Error.

- الحالة الرابع: عدم قبول (رفض) فرضية العدم H_0 وهي في الواقع غير صحيحة، وهذا قرار صحيح.

والجدول (٥-١) التالي يوضح تلك الحالات الأربع:

جدول رقم (٥-١)

الحالات الممكنة حول القرار الذي يتوصل إليه الباحث حول اختبار فرضية العدم

القرار الإحصائي / H_0 في الواقع	رفض H_0	عدم رفض H_0
H_0 صحيحة	قرار خاطئ (خطأ من النوع الأول)	قرار صحيح
H_0 غير صحيحة	قرار صحيح (قوة الاختبار الإحصائي)	قرار خاطئ (خطأ من النوع الثاني)

إن احتمـــال الوقـــوع في الخطأ من النوع الأول، أي احتمـــال رفض فرضية العدم H_0 وهي في الواقع صحيحة يســـمى في الإحصاء بمستوى المعنوية Level of Significance ويرمز له بـ α (تنطق ألفا). كما أن احتمال الوقوع في الخطأ من النوع الثاني، أي احتمال عـــدم رفض (قبول) فرضية العدم H_0 وهي في الواقع غيـــر صحيحة يرمز له بالرمز β (ينطـــق بيتا). هذه الرموز α وβ تمثل قيم احتمالية تتـــراوح قيمها بين الصفر والواحد الصحيح، وهما مرتبطتان عكسياً، بمعنى أن النقصان في احتمالية الوقوع في الخطأ من النوع الأول α يتبعه زيادة في احتمالية الوقوع في الخطأ من النوع الثاني β.

مستوى المعنوية (الدلالة) الاسمي Nominal Significance Level (α):

ويعرف على أنه الحد الأقصى لاحتمال الوقوع في الخطأ من النوع الأول عند اختبار فرضية العدم. ويتم تحديد قيمة مستوى المعنوية من قبل الباحث في المراحل الأولى من إجراء الاختبار الإحصائي. ومن القيم المتعارف عليها والأكثر استخداماً من قبل الباحثين لمستوى المعنوية α هي ٥٪، ١٪ أو ١٠٪. فمثلاً لو تم اختيار $\alpha = 0.05$ فإن ذلك يعني أن عـــدد المرات المحتملة للوقوع في خطأ من النوع الأول (أي عدم قبول فرضية العدم وهي في الواقع صحيحة) هو خمس مرات في كل مائة مرة. بمعنى آخر أنه لو تم سحب مائة عينة عشـــوائية تحت نفس الظروف في كل مرة (أي أن أحجام العينات متساوية وبنفس أسلوب المعاينة ومن نفس المجتمع) لاختبار فرضية عدم معينة فإن عدد القرارت الخاطئة المحتملة حيال فرضية العدم تلك هو خمسة مقابل ٩٥ قراراً صحيحاً.

مستوى المعنوية الفعلي (القيمة الاحتمالية المحسوبة) P-value:

وهي أكبر قيمة لمستوى المعنوية α التي يتم عندها رفض فرضية العدم، وتحسب بناء على قيمة إحصاءة الاختبار المحسوبة من العينة، وبناء على اتجاه الفرضية البديلة وباستخدام توزيع المعاينة لإحصاءة الاختبار تحت فرضية العدم ويرمز لها بالرمز P-value. ويتم مقارنتها بمستوى المعنوية α المحددة سلفاً من قبل الباحث لاتخاذ قرار حول فرضية العدم بالقبول أو عدم القبول. إن برامج التحليل الإحصائي كالـ SPSS وSAS وغيرها تقوم بحساب قيمة الـ P-value لأي اختبار إحصائي يقوم الباحث بتنفيذه من خلال تلك البرامج، وهذا بدوره يسهل المهمة كثيراً حيث إن كل ما على الباحث عمله بعد تحديد الاختبار الإحصائي المناسب للفرضية محل الدراسة هو تنفيذ الاختبار من خلال البرنامج، ومن ثم استخلاص قيمة الـ p-value (تسمى Sig. في الـ SPSS) ومقارنتها بـ α لاتخاذ القرار حول فرضية العدم.

إضاءة إحصائية حول الأخطاء في اختبارات الفرضيات الإحصائية:

يمكن للباحث اختيار قيم أخرى لمستوى المعنوية α حسب مبرراته وحسب طبيعة البحث أو الدراسة، كما أن اختيار هذه القيم يخضع للأثر المترتب في حال الوقوع في الخطأ من النوع الأول أو الثاني. فمثلاً قد تكون الآثار السلبية المترتبة على توصل الباحث إلى أن التدريب لا يؤدي إلى زيادة الإنتاجية، وهو في الحقيقة يؤدي إلى زيادتها (عدم رفض فرضية العدم وهي في الواقع خاطئة: خطأ من النوع الثاني) أكبر من توصله إلى أن التدريب يؤدي إلى زيادة الإنتاجية، وهو في الحقيقة لا يؤدي إلى زيادتها (رفض فرضية العدم وهي في الواقع صحيحة: خطأ من النوع الأول)، وذلك لأن المنظمة لن تقوم بتدريب موظفيها، وهذا بدوره يؤدي إلى عدم مقدرتها على زيادة الإنتاج، وهذا قد يترتب عليه هدر للوقت المستهلك في الإنجاز وعدم المقدرة على تقليص التكاليف وزيادة الأرباح.

قوة الاختبار الإحصائي Power of Statistical Test:

هي قدرة الاختبار على رفض فرضية العدم عندما تكون في الواقع غير صحيحة. وقوة الاختبار ترتبط عكسياً باحتمال الوقوع في الخطأ من النوع الثاني β من خلال العلاقة:

$$\text{Power} = 1-\beta$$

أي أن خفض مستوى β يؤدي إلى زيادة قوة الاختبار. وفي المقابل فإن قوة الاختبار الإحصائي ترتبط طردياً بمستوى المعنوية α، وهذا يقود إلى استنتاج أن العلاقة بين الخطأين α وβ عكسية. أي أن الزيادة /النقص في أحدهما تؤدي إلى النقص/الزيادة في الآخر، انظر الشكل (٥-١). ومما لا شك فيه أن الباحث يرغب في خفض احتمالية الوقوع في أحد الخطأين الأول والثاني (أي تقليص قيمتي α وβ)، وفي الوقت نفسه إلى رفع نسبة قوة الاختبار الإحصائي، وهذا يحتاج إلى موازنة في اختيار تلك القيم ويتم ذلك من خلال تثبيت الخطأ من النوع الأول عند مستوى منخفض نسبياً حسب طبيعة الدراسة والعمل على تخفيض مستوى الخطأ من النوع الثاني إلى أقل حد ممكن كأن يتم زيادة حجم العينة مثلاً. إن تحديد قيمة α والتحكم فيها يعود للباحث، أما قيمة β ومن ثم قوة الاختبار فيعود إلى عدة عوامل سيتم مناقشتها لاحقاً. ومن قيم مستوى المعنوية α المتعارف عليها والمقبولة عملياً بين الباحثين والتي تساعد في تحقيق نوع من التوازن بين الخطأين الأول والثاني هي، ٠,١٠ ، ٠,٠٥ ، ٠,٠١ = α.

حجم التأثير Effect Size (ES):

يعرف حجم التأثير على أنه مؤشر أو مقياس إحصائي لقياس حجم العلاقة بين المتغيرات أو الفروق بين المجموعات. وكما هو معلوم في اختبار الفرضيات (أو اختبارات المعنوية) فإنه في حال رفض فرضية العدم فإن كل ما نستطيع استنتاجه هو وجود علاقة ارتباط معنوية بين المتغيرات أو فروق بين المجموعات، ولا نستطيع من خلالها تقدير قوة الارتباط أو حجم الفروق بين المجموعات. وقد قدم Cohen (1992) تقديراً لحجم التأثير لمجموعة من الاختبارات الإحصائية الأكثر استخداماً، كما وضع تصنيفاً لحجم التأثير من حيث كونه صغيراً، أو متوسطاً، أو كبيراً يختلف باختلاف الاختبار الإحصائي. فمثلاً في حالة الانحدار الخطي يتم استخدام R^2 (معامل التحديد) أو $f^2 = \frac{R^2}{1 - R^2}$ كمؤشر لحجم التأثير ويكون حجم التأثير صغيراً، أو متوسطاً، أو كبيراً للقيم ٠,٢٦ ، ٠,١٣ ، ٠,٠٢ = R^2 على التوالي أو ٠,٣٥ ، ٠,١٥ ، ٠,٠٢ = f^2 على التوالي. وبوجه عام فإن حجم التأثير يعتبر صغيراً إذا كان ٠,٢٠ ≥ ES ومتوسطاً إذا كان ES = ٠,٥٠ وكبيراً ES ≤ ٠,٨٠.

أهمية حجم التأثير:

- يساعد في تحديد المعنوية (أو الأهمية) العملية Practical Significance للمعنوية الإحصائية Statistical Significnace. فقد تكون العلاقة بين المتغيرات معنوية إحصائياً ولكنها غير معنوية عملياً. وقد سبقت الإشارة إلى أن حجم العينة الكبير جداً قد يؤدي إلى استنتاج علاقة أو فروق معنوية إحصائياً، حتى إن كانت غير معنوية أو غير مهمة عملياً، لذا فإنه يمكن تحديد أو تقدير حجم تلك العلاقة أو الفروق بين المجموعات من خلال مؤشر حجم التأثير.
- يمكن أن يستخدم في حساب حجم العينة المطلوب لأي دراسة. إن كلاً من الخصائص الإحصائية التالية: حجم العينة n، حجم التأثير ES، مستوى المعنوية α، وقوة الاختبار الإحصائية Statisitcal Power تعتبر دالة في الأخرى، أي أنه بمعرفة قيم أي ثلاثة من تلك الخصائص يتم تقدير قيمة الخاصية الرابعة. فمثلاً يعتبر حجم العينة n دالة في الخصائص الأخرى، فعند تحديد قيم حجم التأثير وقوة الاختبار الإحصائي ومستوى المعنوية يتم حساب حجم العينة المقابل.
- يعتبر أساساً لتحليل الميتا Meta-analysis، حيث يتم حساب أو تحديد حجم التأثير لكل دراسة من الدراسات المتماثلة في الهدف الداخلة في تحليل الميتا، ومن ثم يتم جمعها بطريقة رياضية معينة ليتم الوصول إلى استنتاج عام حول العلاقة بين المتغيرات أو الفروق بين المجموعات.

العوامل المؤثرة في قوة الاختبار الإحصائي:

من أهم العوامل المؤثرة في قوة الاختبار الإحصائي ما يلي:

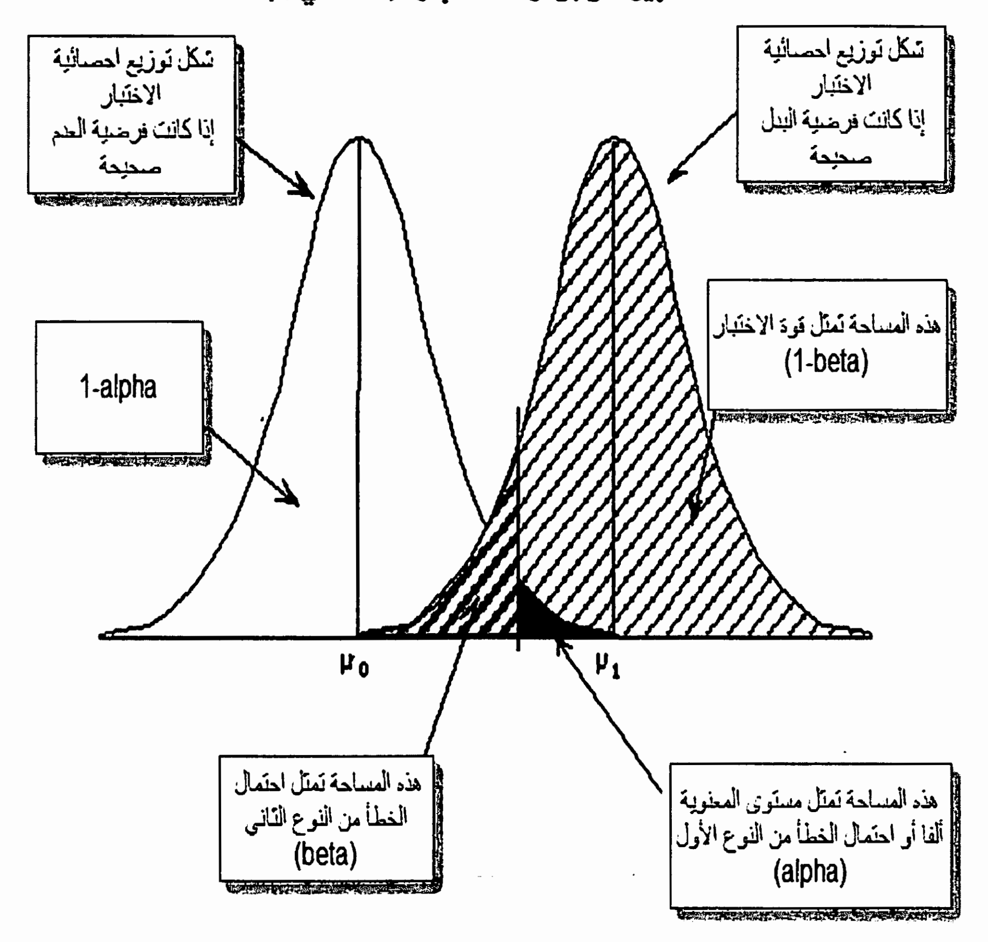
- حجم العينة: فكلما زاد حجم عينة الدراسة زادت مقدرة الاختبار على رفض فرضية العدم. ولكن ينبغي التنويه إلى أن تجاوز العينة حجماً معيناً قد يؤدي إلى تضخيم حجم الخطأ من النوع الأول، كما أن زيادة حجم العينة قد يكون مكلفاً مادياً وزمنياً. وقد ورد عن (زايد، ٢٠٠٧، ٥٩٢) أنه ينبغي اختيار حجم العينة بحيث تكون التكلفه الكلية عند مستواها الأدنى، وذلك وفقاً للصيغة التالية:

التكلفة الكلية = احتمال الخطأ من النوع الأول × تكلفة الخطأ من النوع الأول + احتمال الخطأ من النوع الثاني × تكلفة الخطأ من النوع الثاني + تكلفة المعاينة أو التجربة.

- مستوى المعنوية α: فزيادة قيمة α يؤدي إلى زيادة قوة الاختبار.
- حجم التأثير.
- نـــوع الاختبار فيما إذا كان بطرف (موجـــه) أو بطرفين (غير موجه): فالاختبار ذو الطرف الواحد له قوة أكبر من الاختبار ذي الطرفين.

شكل رقم (٥-١)

العلاقة بين α و β وقوة الاختبار الإحصائي $(1 - \beta)$



الاختبارات الإحصائية المعلمية (البارامترية) Parametric Statistical Tests:

عبارة عن أدوات أو أساليب تتطلب تحقيق افتراضات Assumptions حول معالم ونوع التوزيع الاحتمالي للظاهرة أوالمتغير في المجتمع الذي سحبت منه عينة الدراسة، كأن يكون توزيع الظاهرة أو المتغير في المجتمع طبيعياً (اعتدالياً) Normal، وتجانس التباين بين المجموعات وخطية العلاقة بين المتغيرات. كما ينبغي أن يكون حجم العينة المستخدم في حالة الاختبارات المعلمية كافياً، فقد ذكر سالكايند Salkind (2004) أن غالبية الباحثين اقترحوا ألا يقل حجم العينة عن ٣٠. كما تمت التوصية بأن لا يقل الحد الأدنى للعينة عن ١٠ لكل مجموعة في الدراسة (Warner, 2008). وسيتم تناول أشهر الاختبارات المعلمية وأكثرها استخداماً بالتفصيل لاحقاً.

الاختبارات الإحصائية اللامعلمية (اللابارامترية) Non-Parametric Statistical Tests:

أحياناً تسمى باختبارات التوزيعات الحرة Distribution-Free Tests وهي عبارة عن أدوات أو أساليب إحصائية لا تتطلب تحقيق تلك الافتراضات المطلوب تحقيقها في الاختبارات المعلمية، فهي مثلاً لا تتطلب أن يكون توزيع بيانات المتغير (ات) محل الدراسة طبيعياً، ولا تتطلب عينات كبيرة الحجم. ويتم اللجوء إليها كبديل للاختبارات المعلمية في حال عدم تحقيق بعض من الافتراضات المطلوب تحقيقها للاختبارات المعلمية (Conover, 1999, 113-116). وسيتم تناول أشهر الاختبارات اللامعلمية وأكثرها استخداماً بالتفصيل لاحقاً.

مقارنة بين الاختبارات الإحصائية المعلمية واللامعلمية:

لقد أوضح كونوفر Conover (1999) الظروف والحالات التي تجعلنا نلجأ إلى الاختبارات اللامعلمية كبديل لنظيراتها المعلمية. والجدول (٥-٢) يوضح أبرز أوجه التشابه والاختلاف بين الاختبارات المعلمية واللامعلمية والتي ستسهم في مساعدة وإرشاد الباحث في الاختيار بين الأنسب من البديلين المعلمي واللامعلمي لإجراء الاختبار الإحصائي المطلوب.

العوامل المساعدة في اختيار الاختبار الإحصائي المناسب:

هناك عدة معايير ينبغي على الباحث أو محلل البيانات الإحصائية أخذها بالاعتبار لمساعدته في اختيار الأداة أو الاختبار الإحصائي المناسب في التحليل، ومن أبرزها وأهمها ما يلي:

- الهدف من التحليل.
- هل الهدف مقارنة متوسطات مجموعات، ارتباط بين متغيرات، تنبؤ، ... إلخ.
- عدد المتغيرات المستقلة والتابعة.
- نوع المتغيرات المستقلة والتابعة (متغيرات كمية أو نوعية).
- مستوى قياس المتغيرات المستقلة والتابعة (اسمي، رتبي، فتري أو نسبي).
- مدى تحقق شـــروط التوزيع الطبيعـــي للمتغير التابع، تجانس التباين، اســتقلالية البيانات، خطية العلاقة بين المتغيرات (٨، ٢٠٠٦، Ho).
- حجم العينة أو العينات.

جدول رقم (٥-٢)
مقارنة بين الاختبارات المعلمية واللامعلمية

مقارنة بين الاختبارات المعلمية واللامعلمية	
الاختبارات المعلمية	**الاختبارات اللامعلمية**
تشـــترط توافر معلومات عن نـــوع توزيع بيانات متغير الدراسة التابع كأن يكون توزيعاً طبيعياً.	لا تشترط توافر معلومات عن نوع توزيع بيانات متغير الدراســـة التابع وإنمـــا في حالات تكتفي بأن يكون توزيع المتغير التابع متماثلاً.
تشترط أن تكون عينة/عينات الدراسة عشوائية.	تشترط أن تكون عينة/عينات الدراسة عشوائية.
تشترط أن يكون المتغير التابع فئوياً أو نسبياً.	لا تشترط أن يكون المتغير التابع فئوياً أو نسبياً، أي تستخدم عندما يكون المتغير التابع رتبياً أو اسمياً.
تناسب العينات الكبيرة الحجم.	تناسب العينات الصغيرة الحجم.
تعتمد على القيم الفعلية في حساباتها.	تعتمد على الرتب Ranks في حساباتها أو العد.
إذا كانـــت أطراف التوزيـــع الاحتمالي للبيانات أخف (أقصر) lighter tails من التوزيع الطبيعي – كالتوزيـــع المنتظـــم مثلاً – فـــإن الاختبارات المعلميـــة التي تعتمد على التوزيع الطبيعي تكون لها قوة مساوية أو أكبر من نظيراتها اللامعلمية التي تعتمد على الرتب. وفي الواقع العملي فإن بيانات اســـتطلاعات الرأي openion surveys والتي تســـتخدم مقياس ليكرت الخماســـي أو السباعي على سبيل المثال تعتبر مثالاً للتوزيعات ذات الأطراف الخفيفة (أو القصيرة).	إذا كانت أطراف التوزيع الاحتمالي للبيانات أثقل (أطـــول) Heavier tails مـــن التوزيع الطبيعي- كتوزيع مربع كاي مثلاً – فإن الاختبارات المعلمية التي تعتمد على التوزيع الطبيعي في هذه الحالة تكون لها قوة أقل من نظيراتها اللامعلمية. وفي الواقع العملي تعتبر البيانات التي تحتوي بيانات شـــاذة مثالاً علـــى التوزيعات التـــي لها أطراف أطـــول من التوزيع الطبيعـــي. وفي هذه الحالة ينبغي اســـتخدام الاختبارات اللامعلمية لأن لها قوة أكبر من نظيراتها المعلمية.

تابع - جدول رقم (٥-٢).

مقارنة بين الاختبارات المعلمية واللامعلمية	
الاختبارات المعلمية	الاختبارات الامعلمية
قوة الاختبار الإحصائي تعتمد على شكل توزيع بيانات المجتمع.	قوة الاختبار الإحصائي لا تعتمد على شكل توزيع بيانات المجتمع.
أكثر صعوبة في استخدامها وتسغرق وقتاً أطول في تنفيذها يدوياً.	أسهل في استخدامها وتستغرق وقتاً أقل في تنفيذها يدوياً.

المصدر: من إعداد المؤلف.

خطوات إجراء اختبار الفرضية الإحصائية Statistical Hypothesis Testing:

لقد تم الإشارة إلى أن هناك نوعين من الفرضيات الإحصائية: فرضية العدم والفرضية البديلة، وينبغي على الباحث قبول إحداهما وعدم قبول (رفض) الأخرى، وذلك بالاعتماد على بيانات العينة والاختبار الإحصائي المناسب. إن الاختبارات الإحصائية يمكن تقسيمها إلى الاختبارات المعلمية والاختبارات اللامعلمية. وبغض النظر عن كون الاختبار معلمياً أو لامعلمياً فإن الخطوات المتبعة في تنفيذ تلك الاختبارات هي كالتالي:

١- صياغة فرضيتي العدم H_0 والبديلة H_1.

٢- تحديد مستوى المعنوية أو الدلالة الاسمي α.

٣- تحديد الاختبار الإحصائي المناسب. وقد سبق توضيح العوامل المساعدة في اختيار الاختبار الإحصائي المناسب، وبناء على ذلك يتم تحديد إحصاءة الاختبار Test Statistic وتحديد توزيعها الاحتمالي.

٤- حساب إحصائية الاختبار ومن ثم مستوى المعنوية الفعلي (القيمة الاحتمالية P-value) المرافقة لها.

٥- مقارنة القيمتين الاحتماليتين α وP-value لاتخاذ قرار حيال فرضية العدم مبني على بيانات عينة الدراسة فإذا كانت:

- $\alpha \leq$ P-value فإن الباحث يرفض فرضية العدم.
- $\alpha >$ P-value فإن الباحث لا يرفض («يقبل») فرضية العدم.

٦- اتخاذ القرار وتفسير النتائج ضمن سياق موضوع البحث حيث يعمل الباحث على توثيق استنتاجاته حول فرضية بحثه في ظل معطيات بيانات عينة الدراسة.

ملاحظة: إذا كان الاختبار بطرف واحد فإنه يتم قسمة قيمة P-value التي يتم الحصول عليها من مخرجات التحليل الإحصائي في SPSS على ٢ ومن ثم مقارنة الناتج مع α، ويتم الرفض أو عدم الرفض بنفس الطريقة السابقة في الخطوة رقم (٥).

إضاءة إحصائية حول معنى «معنوي» أو «معنوي إحصائياً»:

سنرى لاحقاً كلمة «معنوي Significant» أو عبارة «معنوي إحصائي Statistical Significance». وهي تعني أن الاختلاف (الفرق) بين المجموعات أو الارتباط بين المتغيرات حقيقي وليس لمحض الصدفة. أي أن بيانات العينة تؤيد بأن الفرق بين متوسطي مجموعتين أو أكثر (مجتمعين أو أكثر) أو الارتباط بين المتغيرات هو فرق أو ارتباط حقيقي وليس بسبب العينة العشوائية المختارة.

إضاءات إحصائية حول بعض الاعتبارات الهامة قبل البدء في التحليل الإحصائي للبيانات:

- يجب أن يكون المقياس (أو أداة جمع البيانات كالاستبانة) صادقاً وثابتاً، وإلا فإنه لا يمكن الاعتماد على البيانات التي جمعت بواسطة تلك الأداة، فما بني على باطل فهو باطل.
- يجب أن تكون عينة الدراسة عينة ممثلة تمثيلاً جيداً لمجتمع الدراسة، وذلك باشتمالها على كل أو معظم صفات وخصائص المجتمع الإحصائي حسب أهداف وتساؤلات الدراسة.
- ينبغي أن يكون حجم العينة مناسباً.
- بعد جمع البيانات، ينبغي استكشافها وتنقيحها وتهيئتها للتحليل الإحصائي، حيث يمكن الاستعانة بأحد برامج التحليل الإحصائي كـ SPSS أو غيره. وينبغي الإشارة إلى أن الجزء والمجهود الأكبر من التحليل الإحصائي للبيانات يتركز في استكشاف وفلترة البيانات وتهيئتها لحساب المؤشرات الإحصائية واختبارات الفرضيات وبناء النماذج الرياضية. فإذا تم تهيئة البيانات وتنقيحها وأصبحت جاهزة للتحليل فإن عملية حساب المؤشرات واختبارات الفرضيات يمكن تنفيذها بكل يسر وسهولة باستخدام أحد برامج التحليل الإحصائي، حيث إن كل ما على الباحث بعد

ذلــك هو تحديد الأســلوب الإحصائي المناســب للتحليل ومن ثــم تنفيذه من خلال تحديد مجموعة من الخطوات البســيطة ويترك للبرنامج تنفيذ الحســابات المعقدة والرســومات البيانية المطلوبة ليتبقى للباحث فقط قراءة وترجمة وتفسير مخرجات التحليل. ولا ننسى بأن برامج التحليل هي كغيرها من البرامج الحاسوبية الأخرى، فهي لا ترشــد الباحث إلى اختيار الأسلوب الإحصائي المناسب للتحليل، وإنما تنفذ آلياً الأوامر التي يحددها المستخدم. وتكمن أهمية استكشاف البيانات قبل تحليلها إحصائياً بشــكل نهائي واعتماد نتائج التحليل في تجنــب الحصول على نتائج غير صحيحة ودقيقة للتحليل الإحصائي. كالوقوع في أحد أو كلا الخطأين الإحصائيين من النوع الأول والثاني، أو إعطاء مؤشــرات إحصائية غير صحيحة، نشوء علاقات ارتباط غير صحيحة، ... إلخ.

وفيما يلي أهم التســاؤلات والخطوط الإرشادية التي ينبغي إجراؤها قبل الإجابة عن تساؤلات وفرضيات الدراسة:

- هل يوجد قيم شــاذة؟ هل تلك القيم الشــاذة هي نتيجة تســجيل خاطئ للبيانات؟ أم أنهــا قيم فعليــة؟ وكيف يتم التعامل معها في كلتا الحالتــين؟ فإذا كانت نتيجة تســجيل أو إدخال خاطــئ في البيانــات أو مأخوذة من مفردة ليســت من مجتمع الدراسة فإنه يتم تعديل الخطأ ومعالجته بالحصول على القيمة الصحيحة أو حذفها من البيانات. أما إذا كانت القيمة الشاذة قيمة فعلية تمثل قياس مفردة من مفردات العينــة التي تنتمي إلى مجتمع الدراســة، فإنه يتم النظر في مدى تأثير تلك القيمة على نتائج التحليل، فإذا كانت لا تؤثر على نتائج التحليل فإنه يتم تجاهلها واعتبارها قيمة عادية كبقية القيم الأخرى في بيانات الدراســة، أما إذا كانت مؤثرة على نتائج التحليل فإنه ينبغي معالجة الموقف إما باســتخدام أساليب إحصائية مقاومة لتأثير القيم الشاذة أو إجراء تحويلات خطية معينة أو إجراء التحليل الإحصائي مع وجود القيم الشاذة وبدونها ومقارنة النتائج وتوضيح الاختلاف.
- هل يوجد قيم مفقودة؟ كم نسبة القيم المفقودة بالنسبة لحجم العينة وكذلك بالنسبة لعدد المتغيرات؟ هل سيتم حذف القيم المفقودة - إما بحذف كل بيانات (متغيرات) المستجيب (الحالة أو المفردة) الذي لديه قيم مفقودة في متغير واحد على الاقل case-wise deletion - وهذا هو الوضع الافتراضي لكثير من برامج التحليل الإحصائي - أوبحذف المتغير الذي به قيم مفقودة فقط والاحتفاظ بالمتغيرات التي ليس بها قيم مفقودة للمســتجيب variable-wise deletion؟ أم من الضروري التعويض عن هذه

القيم المفقودة؟ هل توزيع القيم المفقودة عشوائياً (أي لا يرتبط فقدانها بخصائص أفراد العينة) أم غير عشوائي (أي يرتبط بخصائص أفراد العينة)؟ وما هي الطريقة المناسبة للتعويض إذا دعت الحاجة إلى ذلك؟ وينبغي التنويه إلى أن وجود البيانات المفقودة عند إجراء أساليب التحليل الإحصائي المتعددة المتغيرات يكون أكثر تأثيراً وخطورة - ولاسيما في ظل وجود عينات صغيرة الحجم - على نتائج التحليل، إذ تزيد فرصة حذف الحالات (المستجيبين أو عناصر العينة)، وهذا يؤدي إلى تقليل حجم العينة بشكل كبير وتقليصه إلى الحد الذي يجعل نتيجة التحليل الإحصائي غير سليمة وغير موثوق بها (Tabachnick & Fidell, 2007). لذا فإن التعويض عن القيم المفقودة Missing Values Imputation في هذه الحالة ضروري، وهناك العديد من الطرق للتعويض عن القيم المفقودة، منها:

- المتوسط الحسابي: يتم التعويض عن القيمة المفقودة في متغير ما بمتوسط القيم غير المفقودة في ذلك المتغير. في هذه الحالة متوسط المتغير بعد التعويض لا يتغير ولكن يقل تباينه، وهذا بلاشك يؤدي إلى نتائج متحيزة خاصة إذا كانت القيمة مفقودة بشكل غير عشوائي.
- التعويض باستخدام الانحدار الخطي: يستخدم الانحدار الخطي للتنبؤ بالقيم المفقودة حيث يتم اعتبار المتغير الذي يحوي قيماً مفقودة كمتغير تابع والمتغيرات الأخرى كمتغيرات مستقلة، ويتم بناء نموذج الانحدار واستخدامه في التنبؤ بالقيم المفقودة.
- التعويض المتعدد Multiple Imputation والتعويض باستخدام أسلوب الاحتمال الأعظم Maximum Likelihood: وهي تعتبر من الطرق الحديثة في التعويض عن القيم المفقودة، وتمتاز بأنها تعطي دقة أعلى في حساب القيم المفقودة وتعطي نتائج غير متحيزة نسبياً.

وتجدر الإشارة إلى أن برامج التحليل الإحصائي الشهيرة مثل SPSS وSAS تدعم طرق التعويض السابق ذكرها.

- هل توزيع المتغيرات الكمية في الدراسة طبيعي (Normally Distributed) أم لا؟
- هل توزيع المتغيرات الكمية في الدراسة متماثل symmetric أم غير متماثل asymmetric؟ وفي حال عدم التماثل (الالتواء)، هل الالتواء كبير أم لا؟ وما الإجراء المناسب في حال الالتواء الشديد؟

إضاءة إحصائية حول استخدام اختبارات المعنوية الإحصائية في حال بيانات الحصر الشامل:

إذا كانت بيانات الدراسة بيانات حصر شامل (أي تم الحصول عليها من جمع مفردات مجتمع الدراسة وليس من عينة من المجتمع) فإنه لا يصح وليس منطقياً إجراء اختبار فرضيات (اختبارات المعنوية الإحصائية) حول العلاقة بين المتغيرات أو الفروق بين المجموعات. والسبب يعود إلى أن الهدف من اختبارات الفرضيات واختبارات المعنوية هو المساعدة في اكتشاف - بدرجة احتمال معينة - حقيقة العلاقة بين المتغيرات أو تأثير متغير على آخر في المجتمع من خلال بيانات العينة. فإذا توفرت لنا بيانات المجتمع كاملة فإننا لا نحتاج إلى إجراء اختبارات الفرضيات للوقوف على حقيقة تلك العلاقة بين المتغيرات، وإنما يمكننا مشاهدتها مباشرة من خلال الأساليب الإحصائية الوصفية. وهناك حالات استثنائية مثل بيانات السلاسل الزمنية، فمثلاً قد يتوفر لدى إحدى شركات تأمين المركبات بيانات جميع عملائها خلال عام. ويريد محلل البيانات لدى تلك الشركة بناء أفضل نموذج سلسلة زمنية لعدد طلبات التعويض الأسبوعية في ذلك العام واستخدامه للتنبؤ بعدد طلبات التعويض الأسبوعية لعدد من الأسابيع في العام الذي يليه. هنا إذا أراد المحلل فقط وصف البيانات في ذلك العام فإنه لا يلزمه إجراء اختبارات فرضيات، وإنما يكتفي بالرسم البياني مثلاً و/أو حساب بعض المؤشرات الوصفية مثل أعلى وأقل عدد من طلبات التعويض، المتوسط الحسابي والانحراف المعياري لعدد تلك الطلبات الأسبوعي أو استخدام تحليل مركبات السلسلة الزمنية والمتوسطات المتحركة. ولكن إذا أراد بناء نموذج سلسلة زمنية للتنبؤ بالقيم المستقبلية في العام الذي يليه باستخدام أساليب متقدمة مثل نماذج الانحدار الذاتي Autoregressive Models فإن عليه إجراء بعض اختبارات الفرضيات للوصول إلى أفضل نموذج، ومن ثم إلى درجة دقة أكبر في التنبؤ.

الفصل السادس
الاختبارات المعلمية للفروق بين المتوسطات
Parametric Tests for Mean Differences

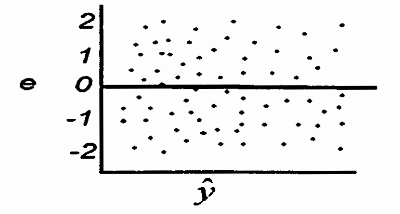
مقدمة:

لقد سـبقت الإشـارة في الفصل الخامس إلى أن الاختبارات وأسـاليب التحليل الإحصائي المعلمية تعتمد على مجموعة من الافتراضات أو الشروط (Assumptions) اللازمة لضمان صحة ودقة نتائج التحليل. لذا ينبغي التأكد من أن الافتراضات اللازمة لاستخدام الاختبار الإحصائي متحققة، وذلك لأن عدم تحقق أحد هذه الشروط يؤدي إلى تقليل أو عدم الثقة في مصداقية نتائج الاختبار، وهذا بدوره سيؤثر في مصداقية الدراسـة أو البحث وعدم الثقة في نتائجه. ومعظم الاختبارات أو أسـاليب التحليل المعلمية تشـترك في مجموعة من تلك الافتراضات، لذا ينبغي على الباحث التأكد من صحة الافتراضات اللازمة للاختبار أو الأسلوب الإحصائي المعلمي محل الدراسة قبل اعتماد نتائـج التحليل. وفيما يلي عرض مختصر لتلـك الافتراضات وكيفية فحص تحققها من عدمه:

- أن تكون وحدات المعاينة مختارة بطريقة عشـوائية (احتمالية)، وذلك تجنباً للتحيز في تقدير معلمات مجتمع الدراسة وضماناً لاستقلالية البيانات بعضها عن بعض.
- التوزيع الطبيعي للبيانات Normality. وللتحقق من أن بيانات المجتمع الذي سحبت منه العينة تتبع التوزيع الطبيعي، يتم اسـتخدام طرق بيانيـة مثل المدرج التكراري، رسم الطبيعية أو الاعتدالية Normality plot، رسم q-q البياني وغير ذلك. كذلك يتم بطرق غير بيانيـة (أي اختبارات إحصائية) لاختبار الاعتدالية لتوزيع البيانات، ومنها على سـبيل المثال لا الحصر اختبار كولمقورف - سـميرنوف Kolmogorov-Smirnov واختبار شابيرو - ويلك Shapiro-Wlik.
- تجانس التباين Homogeniety of Variances. وللتحقق من تسـاوي التباينات بين مجموعات مجتمع الدراسـة، يفضل اسـتخدام اختبار ليفين لتساوي التباينات Levene's Test for Equality of Variances.
- اسـتقلالية البيانات Independence of data. إن التحقق من صحة هذا الشـرط يعتمـد بشـكل كبير على تصميم المعاينة العشـوائية بطريقة تضمـن إلى حد كبير اسـتقلالية قيـم (ظهور أحدها لا يرتبـط بحدوث الآخر) المتغيـر التابع أو الخطأ

العشـــوائي بعضها عن بعض. ولكن عندما يتبع المتغير التابع ترتيباً تسلسلياً طبيعياً تبعاً لمتغير معين كالزمن فإنه يتم عادة استخدام الأشكال الانتشارية للبقايا Residual plots أو الأخطـــاء العشـــوائية Random error plots (ويرمز لها بالرمز e) الناتجة عـــن تقدير النماذج الإحصائيـــة كنماذج تحليل التباين ونمـــاذج تحليل الانحدار مع القيم التنبؤية لقيم المتغير التابع $\hat{y}$. فإذا كانت القيم متناثرة بشـــكل عشـــوائي حول قيمـــة الصفـــر لـ e كما هو موضح في الشـــكل (٦-١) فإنه يمكـــن القول بأن قيم y أو e مســـتقلة. وكذلك يمكن اســـتخدام اختبار دوربن - واتسن Durbin-Watson، كما يمكن اســـتخدام دالة الارتباط الذاتي ACF أو الرسم البياني للارتباط الذاتي Correlogram. إن عـــدم تحقـــق شـــرط الاســـتقلالية يؤثر على نتائـــج اختبارات الفرضيات ويؤدي إلى تضخم الخطأ من النوع الأول α، أي الفشل في رفض فرضية العـــدم وهي خاطئة (Glass & Hopkins, 1995). لذا لابد من اســـتخدام طرق أو أســـاليب إحصائية أخرى كنمـــاذج ARIMA في حالة السلاســـل الزمنية بدلاً من الانحدار الخطي الاعتيادي.

شكل رقم (٦-١)
الشكل الانتشاري للعلاقة بين e و$\hat{y}$



- الخطيـــة Linearity: فـــي النماذج الخطية يجب أن يكـــون المتغير التابع خطياً في معالم النموذج. وفي الانحدار الخطي يمكن تجسيد الخطية في أن تتغير متوسطات التوزيعات الاحتمالية للمتغير التابع بصورة خطية مع التغير في قيم المتغيرات المستقلة. وبوجه عام فإنه ينبغي أن تكون العلاقة بين المتغير التابع والمتغيرات المستقلة التي تم بناؤها من بيانات العينة هي «نفس» العلاقة بين المتغير التابع والمتغيرات المستقلة في المجتمع الذي ســـحبت منه العينة. ويمكن فحص خطية النموذج من خلال الرســـوم البيانية المناســـبة، مثل الشكل الانتشـــاري Scatter plot للعلاقة بين المتغير المستقل والتابع والرســـم البياني للبقايا Residuls plot. إن عدم تحقق هذا الشرط سيؤثر على درجة الدقة في التنبؤ وسيحد من تعميم النتائج.

- مشكلة الازدواج الخطي (الخطية المشتركة المتعددة) Multicollinearity. في تحليل الانحدار (وغيره من الأساليب الإحصائية المتعددة المتغيرات كالتحليل العاملي مثلاً) ينبغي التأكد من عدم وجود هذه المشكلة والتي تعني وجود ارتباط عالٍ بين متغيرين مستقلين أو أكثر. وهناك طرق متعددة يمكن من خلالها التعرف على وجود تلك المشكلة. ومن تلك الطرق استخدام مصفوفة الارتباط الخطي بين المتغيرات والتي توضح درجة الارتبط بين كل متغيرين. كما يمكن استخدام معامل تضخم التباين Variance Inflation Factor(VIF) لتحديد المتغيرات المرتبطة ببعضها ارتباطاً عالياً. إنه ينبغي فحص مشكلة الازدواج الخطي أولاً قبل الشروع في اختبار الفرضيات وبناء النماذج الإحصائية؛ لأنها تؤدي إلى تغيير وحجب ملامح العلاقة بين المتغيرات في الدراسة، ومن ثم لا يمكن الاعتماد على نتائج التحليل الإحصائي في حال وجودها. وقد تطرق آندي فيلد Field (2005) إلى تأثير الازدواج الخطي على تحليل الانحدار الخطي، فقد ذكر بأن وجود مشكلة الازدواج الخطي سيزيد من احتمالية عدم القبول إحصائياً لمتغير مستقل جيد وهام للتنبؤ بالمتغير التابع (أي سيزيد من تضخم الخطأ من النوع الثاني). كما أن ذلك سيؤثر على التعرف على أي من المتغيرات المستقلة المرتبطة ارتباطاً قوياً هو أكثر أهمية بالنسبة للمتغير التابع. بالإضافة إلى ذلك فإن هذه المشكلة ستزيد من حجم تباين معاملات الانحدار ومن ثم يؤثر على دقة التنبؤ. علاوة على ما ذكر فقد تؤدي هذه المشكلة إلى عكس اتجاه العلاقة بين المتغير المستقل والتابع.
- وقد ناقش قلاس وهوبكنز Glass &Hopkins (1995) بعض الحقائق حول مخالفة افتراضات الاعتدالية وتجانس التباينات، والاستقلالية لاختبارات F وt للمتوسطات حيث ذكرا بأن:
  - اختبارات F وt للمتوسطات تعتبر حصينة Robust في حال عدم تحقق افتراض الاعتدالية مادامت درجة عدم الاعتدالية غير شديدة، وأحجام العينات كبيرة، والاختبار ليس في اتجاه واحد. بمعنى أن عدم اعتدالية البيانات لا تؤثر في مصداقية نتائج الاختبار وسيكون لها تأثير طفيف بسيط جداً غير مؤثر على الخطأين من النوع الأول والثاني ومن ثم على قوة الاختبار.
  - كذلك فإن تلك الاختبارات تُعد أيضاً حصينة في حال مخالفة شرط تجانس التباينات بين المجموعات مادامت أحجام العينات في المجموعات متساوية. حيث أن مخالفة تجانس التباينات سيؤثر بشكل طفيف جداً على مستوى الخطأ من النوع الأول وقوة الاختبار.

- ولكنها غير حصينة في حال مخالفة شرط استقلالية البيانات؛ إذ إن عدم تحقق هذا الشرط سيؤدي إلى تضخم الخطأ من النوع الأول بشكل كبير.

وبناء على ما سبق فإنه ينبغي معالجة تلك المخالفات والمشاكل قبل اعتماد نتائج الاختبارات وأساليب التحليل الإحصائي المعلمية أو اللجوء إلى أساليب إحصائية أخرى كالأساليب اللامعلمية.

وفي هذا الفصل سيتم تقديم أكثر الاختبارات الإحصائية المعلمية استخداماً لاختبار الفرضيات حول المتوسطات وشرحها بطريقة عملية تفاعلية باستخدام برنامج SPSS (القحطاني، ١٤٣٢).

أولاً - اختبار t لمتوسط عينة واحدة One Sample t Test:

الهدف من استخدامه:

لاختبار الفرضية حول متوسط مجتمع واحد، وذلك من خلال اختبار فيما إذا كان متوسط العينة يختلف اختلافاً معنوياً (حقيقياً) عن القيمة الافتراضية لمتوسط المجتمع.

متى يستخدم:

عندما يكون تباين المتغير في مجتمع الدراسة غير معلوم وفي الوقت نفسه عندما يكون حجم العينة صغيراً، وتحديداً عندما يكون حجم العينة (n) أقل من ٣٠.

افتراضات (شروط) استخدامه The assumptions:

١- أن تكون بيانات العينة مسحوبة عشوائياً من مجتمع الدراسة ومستقلة بعضها عن بعض.

٢- أن يكون متغير الدراسة متصلاً (من مستوى القياس الفئوي أو النسبي).

٣- أن يكون توزيع متغير الدراسة في مجتمع الدراسة طبيعياً (اعتدالياً).

ملاحظة: علاوة على الافتراضات السابقة، إذا كان تباين المتغير في المجتمع معلوم فإنه يتم استخدام اختبار Z بدلاً من t بغض النظر عن حجم العينة. أيضاً يفضل استخدام اختبار Z إذا كان تباين المتغير في المجتمع غير معلوم، ولكن بشرط أن يكون حجم العينة على الأقل ٣٠. وهذا ينطبق على كل اختبارات t اللاحقة.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

سيتم توضيح كيفية تنفيذ اختبار t وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي:

تطبيق عملي تفاعلي (٦-١):

في إحدى المدارس الثانوية بمدينة ما أراد باحث التحقق مما إذا كان متوسط ذكاء (أو ما يعرف بالـ IQ) طلاب تلك المدرسة يختلف عن متوسط الذكاء العادي والذي يساوي ١٠٠. لذا قام الباحث باختيار ٢٠ طالباً من طلاب الثانوية لتلك المدرسة بطريقة عشوائية وطبق عليهم اختبار الذكاء وكانت النتائج كالتالي:

جدول رقم (٦-١)
مستوى ذكاء طلاب إحدى المدارس الثانوية بمدينة ما

١٠٠	٩٩	١٢٠	١٠١	١٠٥	١٠٦	١١٠	١١١	١٢٠	٩٨
٩٥	١٠٠	١٢١	١١٠	١٠٣	١٠٩	١٠٠	١٢٠	١٠٠	٩٩

وسؤال البحث هو: هل يختلف متوسط ذكاء طلاب تلك المدرسة عن متوسط الذكاء العادي = ١٠٠؟

خطوات الحل:

١- صياغة الفرضيات:

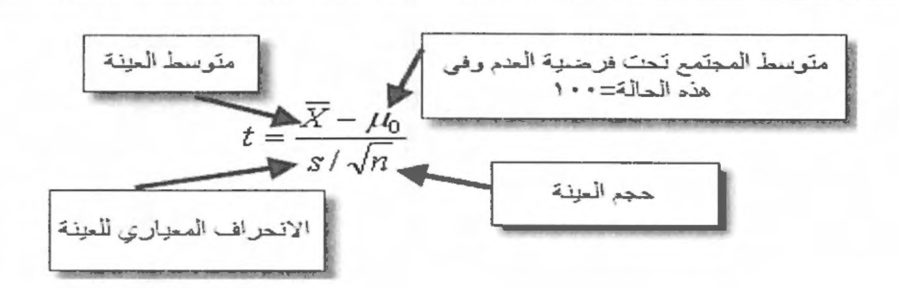
H_0: لا يوجد اختلاف معنوي بين متوسط ذكاء طلاب تلك المدرسة ومتوسط الذكاء العادي والذي يساوي ١٠٠ (أي أن $\mu = \mu_0 = 100$).

H_1: يوجد اختلاف معنوي بين متوسط ذكاء طلاب تلك المدرسة ومتوسط الذكاء العادي ($\mu = \mu_0 \neq 100$).

ونلاحظ: أن الفرضية البديلة هنا ذات اتجاهين، ومن ثم فإن اختبار t في هذه الحالة يعد اختباراً بطرفين Two-tailed test.

٢- تحدد مستوى المعنوية: وليكن $\alpha =$ ٠,٠٥.

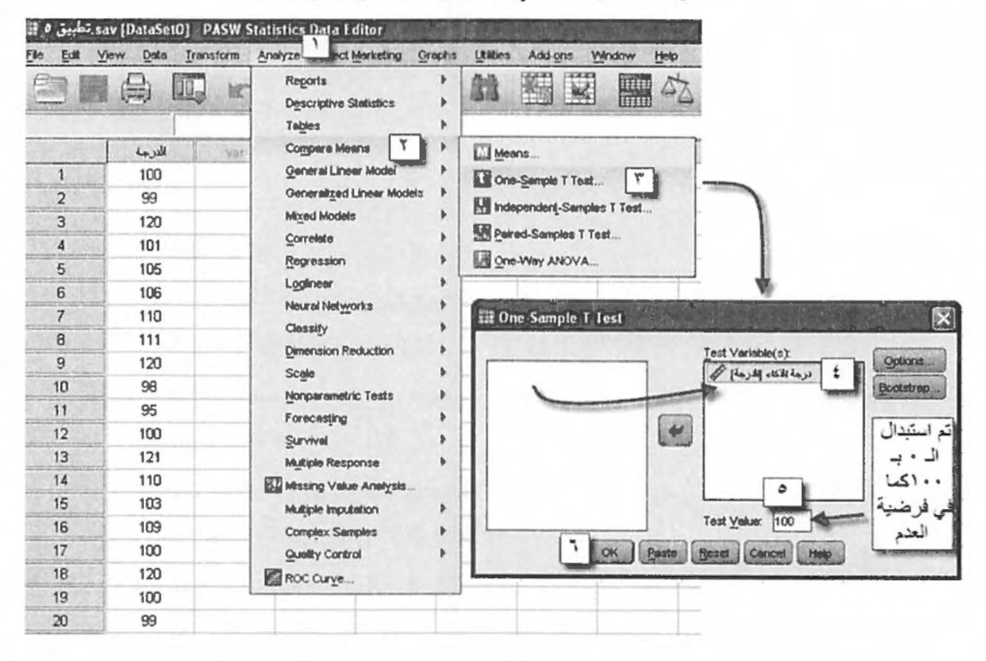
٣- اختيار الاختبار الإحصائي المناسب: الاختبار الإحصائي في هذه الحالة هو t لمتوسط عينة واحدة (One Sample t Test) ومن ثم تأخذ إحصائية الاختبار الشكل التالي:



إحصاءة الاختبار هذه تتبع توزيع t الاحتمالي بدرجات حرية n-1 (أي أن $t \sim t_{(n-1)}$).

٤- الآن قم بادخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومـن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٦ كما هو موضح بالشكل (٦-٢) لتنفيذ الاختبار.

شكل رقم (٦-٢)
خطوات تنفيذ اختبار t لمتوسط عينة واحدة



وبعد التنفيذ سيتم الحصول على النتائج التالية في نافذة المخرجات Output كالتالي:

شكل رقم (٦-٣)
مخرجات تنفيذ اختبار t لمتوسط عينة واحدة

T-Test

مخرجات التحليل

[DataSet0] C:\Documents and Settings\kahtaniss\Desktop\5 الحقيبة\تطبيق1.sav

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
درجة الذكاء	20	106.35	8.393	1.877

p-value

One-Sample Test

	Test Value = 100					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
درجة الذكاء	3.383	19	.003	6.350	2.42	10.28

٥- اتخاذ القرار وتفسير النتائج: بما أن $(٠,٠٥)\ a > (٠,٠٠٣)$ P-value فإن الباحث سيقوم برفض فرضية العدم ومن ثم قبول فرضية البدل. وهذا يقود إلى استنتاج أنــه عند مســتوى معنوية $a = ٠,٠٥$ فــإن بيانات العينة تؤكد أن متوســط الذكاء لطلاب الثانوية بتلك المدرسة (١٠٦,٣٥) يختلف عن متوسط ذكاء الطالب العادي (١٠٠)، (t (١٩) = ٣,٣٨ ، p = ٠,٠٠٣).

تمرين تطبيقي (٦-١):

أراد باحث التحقق مما إذا كان معدل الاســتهلاك الشــهري للأســرة من اللحوم الحمــراء في إحدى المدن يفوق معدل الاســتهلاك الشــهري العام للأســرة (٨) كلغ، لذا قام باختيار عينة عشــوائية مكونة من ١٠ من الأســر من تلك المدينة، وكان معدل استهلاكهم الشهري من اللحوم الحمراء كما يلي:

جدول رقم (٦-٢)

معدل الاستهلاك الشهري للأسرة من اللحوم الحمراء (بالكيلوجرام) في إحدى المدن

الأسرة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
معدل الاستهلاك الشهري (كغم)	٤	٦	٨	٦	١٠	٥	١٢	٨	٧	٩

عند مستوى معنوية $a =$ ٠,٠٥، هل يختلف متوسط الاستهلاك الشهري للأسرة في تلك المحافظة معنوياً عن المتوسط العام للاستهلاك الشهري (٨) كلغم؟

ثانياً - اختبار t لمتوسطي عينتين مرتبطتين Paired-samples t test:

الهدف من استخدامه:

لاختبار الفرضية حول متوسطي مجتمعين مترابطين، وذلك من خلال اختبار ما إذا كان الفرق بين متوسطي العينتين المرتبطتين يختلف اختلافاً معنوياً (حقيقياً) عن الفرق بين متوسطي مجتمعي الدراسة.

شروط استخدامه:

١- العينتان مرتبطتان.

٢- أن تكون العينتان مختارتين عشوائياً من مجتمعي الدراسة.

٣- أن يكون المتغير التابع محل الدراسة متصلاً (فئوي أو نسبي).

٤- أن يكون توزيع الفرق بين قيم المتغير التابع في مجتمعي الدراسة طبيعياً.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

سيتم توضيح كيفية تنفيذ اختبار t لمتوسطي عينتين مرتبطتين وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٦-٢):

أراد باحث دراسة مدى تأثير الدعاية والإعلان للمنتجات المحلية من خلال القنوات الفضائية على طلب المستهلك لتلك المنتجات. لذا قام باختيار عينة عشوائية مكونة من سبعة منتجات محلية ورصد مبيعاتها الأسبوعية بآلاف الريالات قبل الدعاية، ثم قام برصد مبيعاتها الأسبوعية بآلاف الريالات بعد الإعلان لها من خلال القنوات الفضائية، وكانت النتائج كالتالي:

جدول رقم (٦-٣)

المبيعات الأسبوعية لسبعة منتجات محلية (بآلاف الريالات) قبل وبعد الدعاية من خلال القنوات الفضائية

المنتج	١	٢	٣	٤	٥	٦	٧
قبل	٤٤,٢	١٩,٠	٢١,٨	٣٢,٥	١٨,٢	٤٥,٨	٢٢,٥
بعد	٤٨,٥	٢٣,٣	٢٢,٠	٣٤,٤	١٨,١	٥٠,٠	٢٦,٣

والسؤال هل الدعاية للمنتجات تؤدي إلى زيادة مبيعاتها؟

خطوات الحل:

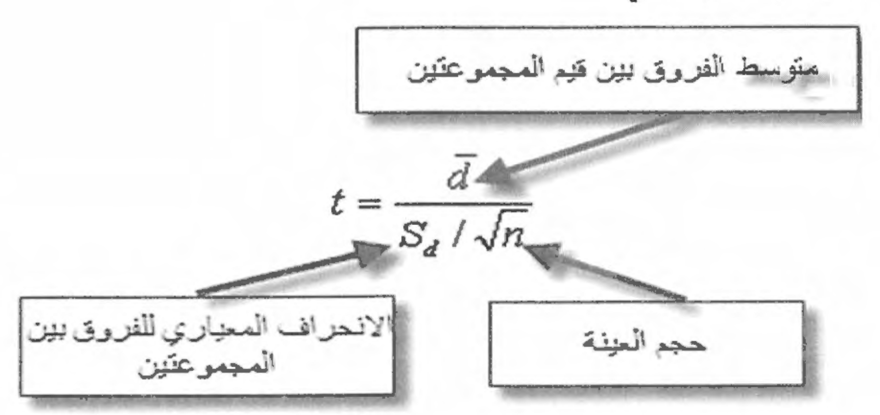
١- صياغة الفرضيات:

H_0: لا يوجد اختلاف معنوي بين متوسطي المبيعات للمنتجات قبل الدعاية وبعدها.

H_1: متوسط المبيعات للمنتجات بعد الدعاية أكبر من متوسطها قبل الدعاية.

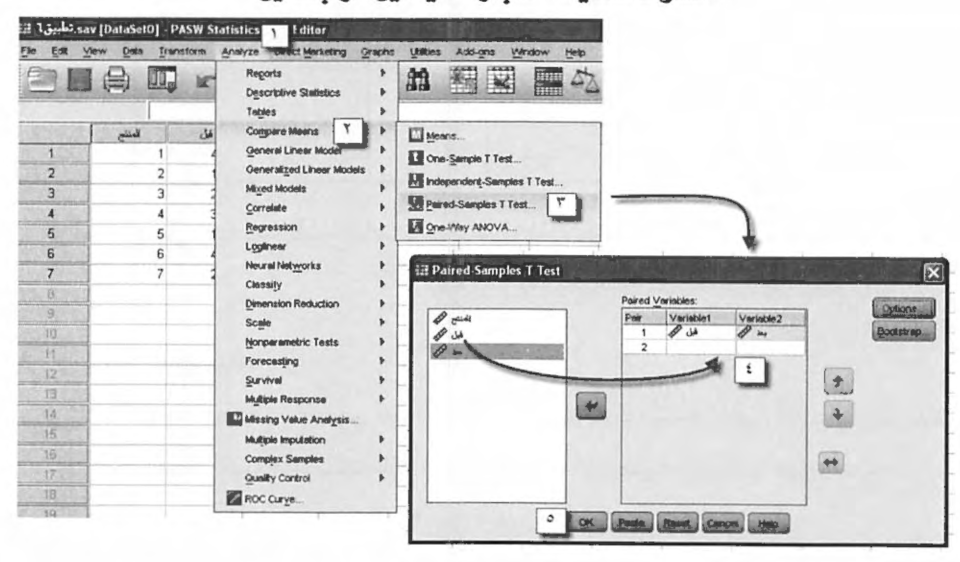
٢- تحديد مستوى المعنوية: وليكن $a =$ ٠,٠٥.

٣- تحديد الاختبار الإحصائي المناسب: الاختبار الإحصائي في هذه الحالة هو t لمتوسطي عينتين مرتبطتين (Paired-samples t Test) ومن ثم تأخذ إحصائية الاختبار الشكل التالي:



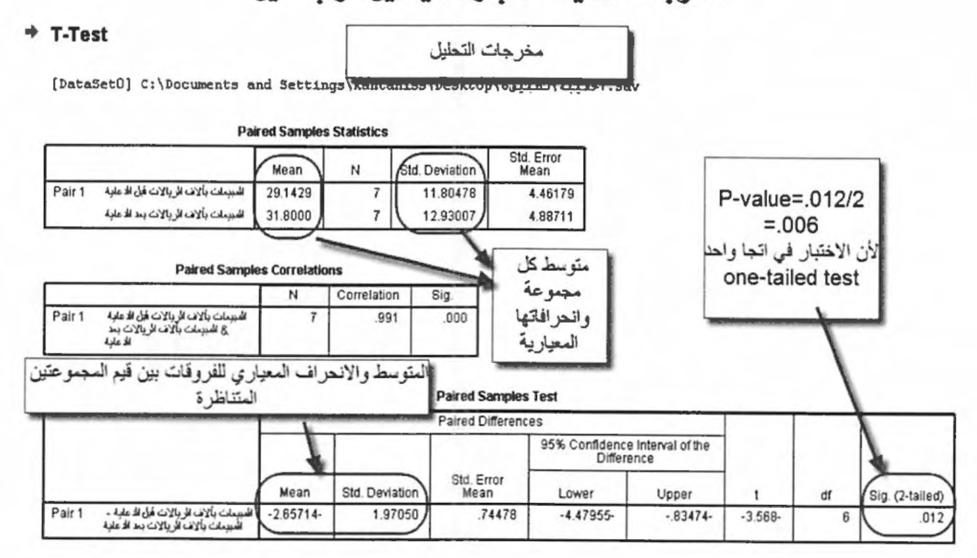
٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٥ كما هو موضح بالشكل (٦-٤) لتنفيذ الاختبار.

شكل رقم (٦-٤)
خطوات تنفيذ اختبار t لعينتين مرتبطتين



وبعد التنفيذ يتم الحصول على النتائج التالية في نافذة المخرجات Output كالتالي:

شكل رقم (٦-٥)
مخرجات تنفيذ اختبار t لعينتين مرتبطتين



T-Test

[DataSet0] C:\Documents and Settings\...

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	المبيعات بآلاف الريالات قبل الدعاية	29.1429	7	11.80478	4.46179
	المبيعات بآلاف الريالات بعد الدعاية	31.8000	7	12.93007	4.88711

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	المبيعات بآلاف الريالات قبل الدعاية & المبيعات بآلاف الريالات بعد الدعاية	7	.991	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	Upper			
Pair 1	المبيعات بآلاف الريالات قبل الدعاية - المبيعات بآلاف الريالات بعد الدعاية	-2.65714-	1.97050	.74478	-4.47955-	-.83474-	-3.568-	6	.012

٥- اتخاذ القرار وتفسير النتائج: بما أن (٠,٠٥) a > (٠,٠٠٦) P-value، فإن الباحث سيقوم برفض فرضية العدم ومن ثم قبول الفرضية البديلة. وهذا يقود إلى استنتاج أنه عند مستوى معنوية ٠,٠٥ = a فإن بيانات العينة تؤكد أن متوسط المبيعات الأسبوعية للمنتجات بعد الدعاية (٣١,٨ ألف) أعلى منه قبل الدعاية (٢٩,١٤ ألف)، (p-value = ٠,٠٠٦ ، t (٦) = -٣,٥٧).

تمرين تطبيقي (٦-٢):

قامت إحدى الشركات بصنع جهاز يساعد في خفض معدل استهلاك السيارة للوقود. وللوقوف على مدى فاعلية الجهاز، نفذت الشركة دراسة باختيار عشر سيارات مختلفة عشوائياً ورصدت المسافة (بالأميال) التي قطعتها كل سيارة للجالون الواحد بدون الجهاز وبعد تركيبه، وكانت النتائج كالتالي:

جدول رقم (٦-٤)
المسافة المقطوعة (بالأميال) بالجهاز وبدونه

السيارة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
بدون الجهاز	٢٤,٩	١٨,٨	٢٧,٧	١٣	١٧,٨	١١,٣	٢٧,٨	٨,٢	٢٣,١	٩,٩
بالجهاز	٢٥,٧	٢٠,٠	٢٨,٤	١٣,٧	١٨,٨	١٢,٥	٢٨,٤	٨,١	٢٣,١	١٠,٤

والمطلوب: عند مستوى معنوية ٠,٠٥ = α وبالاعتماد على بيانات عينة الدراسة، هل يؤدي الجهاز إلى زيادة المسافة المقطوعة بالأميال لكل جالون من الوقود؟

تمرين تطبيقي (٦-٣):

لاختبار تأثير عقار طبي جديد يعمل على خفض مستوى الكوليسترول في الدم في أحد مراكز الأبحاث الطبية، تم إجراء دراسة على عينة عشوائية قوامها ١٠ من الفئران، والجدول التالي يوضح مستوى الكوليسترول في الدم قبل وبعد حقن تلك الفئران بالعقار الجديد:

جدول رقم (٦-٥)
مستوى الكولسترول في الدم قبل وبعد العقار الطبي الجديد

الفأر	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
الكوليسترول بعد العقار الطبي	٢٢٢	٢٣٠	٢٣٦	٢٤٨	١٩٦	٢٤٨	٢١٣	٢١١	١٩٩	٢٤٤
الكوليسترول قبل العقار الطبي	٢٣٥	٢٦٢	٢٣٠	٢٦٧	٢٣٥	٢٦٩	٢٨٩	٢٠٣	٢٢٥	٢٨٧

والمطلوب: عند مستوى معنوية $\alpha = 0.01$، هل تؤيد بيانات العينة بأن العقار الطبي الجديد قد أسهم في تخفيض مستوى الكوليسترول في الدم؟

ثالثاً - اختبار t لمتوسطي عينتين مستقلتين Two Independent Samples t Test:

الهدف من استخدامه:

لاختبار الفرضية حول متوسطي مجتمعين مستقلين، وذلك من خلال اختبار ما إذا كان الفرق بين متوسطي العينتين المستقلتين يختلف اختلافاً معنوياً (حقيقياً) عن الفرق بين متوسطي مجتمعي الدراسة.

متى يستخدم:

عندما يكون تباينا مجتمعي الدراسة اللذان سحبت منهما العينات غير معلومة مسبقاً، وفي نفس الوقت تكون أحجام عينتي الدراسة صغيرة (أقل من ٣٠). أي عندما تكون $n_1, n_2 < 30$ حيث n_1 وn_2 ترمزان لحجمي العينتين المسحوبتين من مجتمعي الدراسة الأول والثاني على التوالي.

شروط استخدامه:

١- العينتان مستقلتان.

٢- أن تكون العينتان مختارتين عشوائياً من مجتمعي الدراسة.

٣- أن يكون المتغير التابع محل الدراسة متصلاً (فئوياً أو نسبياً).

٤- أن يكون توزيع المتغير التابع في مجتمعي الدراسة طبيعياً.

٥- أن يكون تباين المتغير التابع في المجتمع الأول يساوي تباينه في المجتمع الثاني.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

سيتم توضيح كيفية تنفيذ اختبار t لمتوسطي عينتين مستقلتين وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي:

تطبيق عملي تفاعلي (٦-٣):

في دراسة لفريق طبي لمعرفة ما إذا كان هناك فرق في عمر الطفل عند المشي بين الذكور والإناث، تم اختيار عشرة أطفال من كل نوع عشوائياً وتم تدوين عمر المشي

بالشهور لكل منهم كما هو موضح بالجدول أدناه. فهل تؤيد البيانات أن هناك اختلافاً معنوياً بين متوسطي أعمار المشي عند الأطفال الذكور والإناث؟ اجعل $\alpha =$ ٠,٠٥.

جدول رقم (٦-٦)
أعمار الأطفال الذكور والإناث (بالشهور) عند المشي

الذكور	٩,٠	١٠,١	٩,٢	١٠,٢	١٠,٠	١٢,٨	١٣,٤	٩,٧	١٠,٥	١١,١
الإناث	٩,٥	١٢,٣	١٣,٢	١٢,٦	١٣,٤	٩,٦	٩,٨	١٣,٢	١٢,٠	١٠,٢

خطوات الحل:

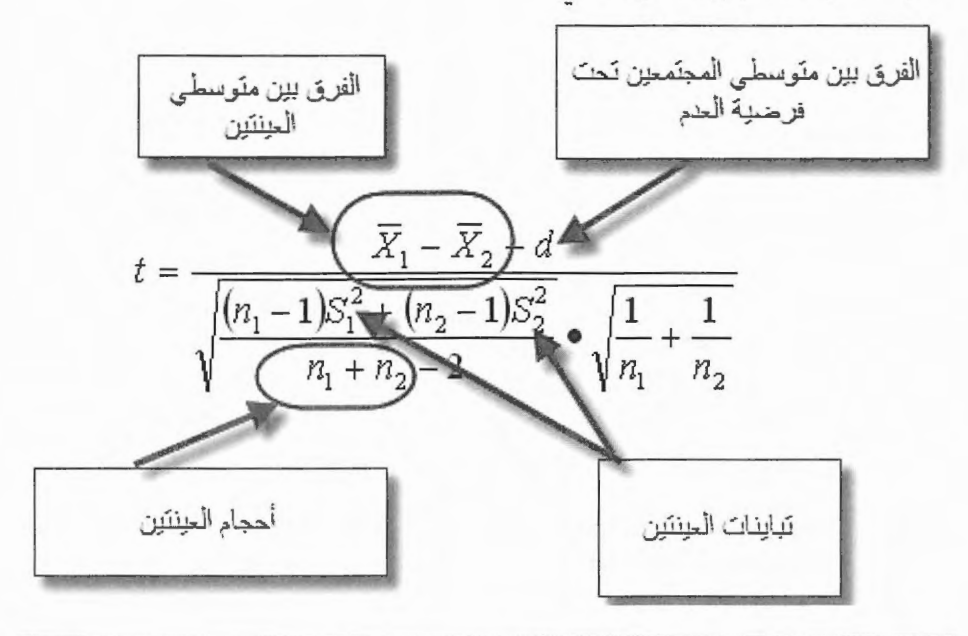
١- صياغة الفرضيات:

H_0: لا يوجد اختلاف معنوي بين متوسطي عمر المشي للذكور والإناث.

H_1: يوجد اختلاف معنوي بين متوسطي عمر المشي للذكور والإناث.

٢- تحديد مستوى المعنوية: وليكن $\alpha =$ ٠,٠٥.

٣- تحديد الاختبار الإحصائي المناسب: الاختبار الإحصائي في هذه الحالة هو t لمتوسطي عينتين مستقلتين (Independent Samples T Test)، ومن ثم تأخذ إحصائية الاختبار الشكل التالي:



$$t = \frac{\bar{X}_1 - \bar{X}_2 - d}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

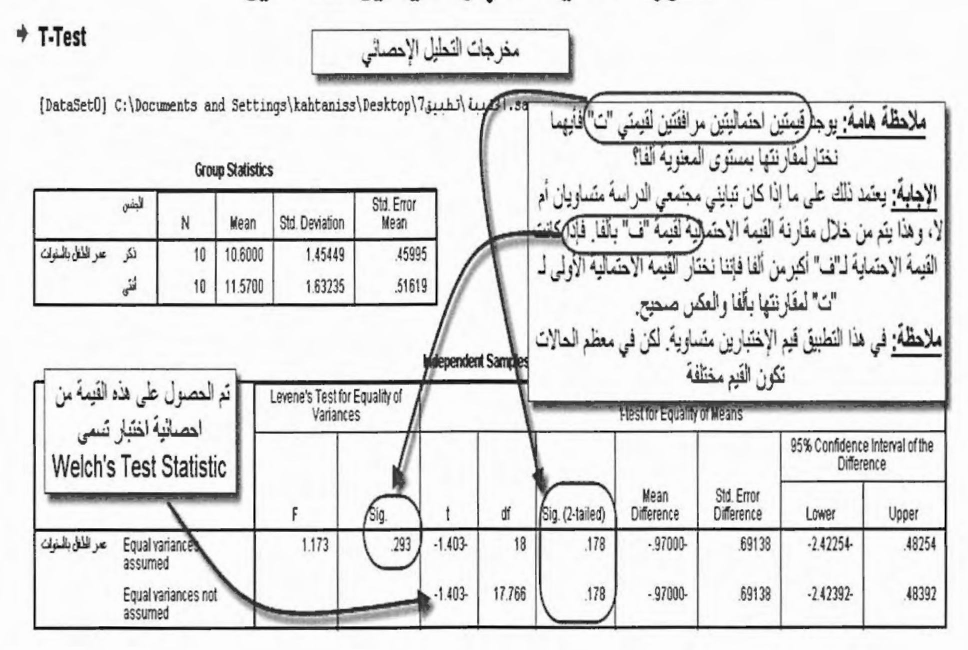
٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومـــن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حســـب الأرقام التسلسلية من ١ إلى ٩ كما هو موضح بالشكل (٦-٦) لتنفيذ الاختبار.

شكل رقم (٦-٦)
خطوات تنفيذ اختبار t لعينتين مستقلتين

وبعد تنفيذ الاختبار يتم الحصول على النتائج التالية في نافذة المخرجات Output كالتالي:

شكل رقم (٦-٧)

مخرجات تنفيذ اختبار t لعينتين مستقلتين



T-Test

[DataSet0] C:\Documents and Settings\kahtaniss\Desktop\7تطبيق\ ...sa

Group Statistics

	الجنس	N	Mean	Std. Deviation	Std. Error Mean
عمر الأطفال بالسنوات	ذكر	10	10.6000	1.45449	.45995
	أنثى	10	11.5700	1.63235	.51619

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
عمر الأطفال بالسنوات	Equal variances assumed	1.173	.293	-1.403-	18	.178	-.97000-	.69138	-2.42254-	.48254
	Equal variances not assumed			-1.403-	17.766	.178	-.97000-	.69138	-2.42392-	.48392

٥- اتخاذ القرار وتفسير النتائج: بما أن α (٠,٠٥) < (٠,١٨) P-value فإن الباحث لا يرفض فرضية العدم. وهذا يقود إلى استنتاج أنه عند مستوى معنوية α = ٠,٠٥ فإنه ليس هناك دليل كاف على أن متوسط عمر المشي للأطفال الذكور (١٠,٦) يختلف معنوياً عن متوسط عمر المشي للأطفال الإناث (١١,٥٧)، (t (٩،٩) = −١,٤٠ ، p-value = ٠,١٨).

إضاءة إحصائية حول اختبار t لمتوسطي عينتين مستقلتين:

نلاحظ عند تنفيذ اختبار t لمتوسطي عينتين مستقلتين باستخدام SPSS أو SAS فإنه ينتج لدينا قيمتين لهذا الاختبار. وكل ما يحتاجه الباحث هو أحد تلك القيمتين؟ فلماذا قيمتان وأيهما يختار الباحث؟ إن هاتين القيمتين تعتمدان على ما إذا كان تباين المتغير التابع في المجتمع الأول (σ^2_2) يساوي تباينه في المجتمع الثاني (σ^2_1) أو لا يساويه. فإذا كان تباينا المجتمعين متساويين فإنه يتم استخدام إحصائية الاختبار التالية:

$$t=\frac{(\bar{X}_1-\bar{X}_2)-d}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}\times\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

وهذه الإحصائية تتبع توزيع t الاحتمالي بدرجات حرية degrees of freedom (df) تساوي 2 - $df = n_1 + n_2 - 2$، أي أن $t \sim t_{(n+n2-2)}$.

أما إذا كان تباينا المجتمعين غير متساويين فإنه يتم استخدام إحصائية الاختبار التالية:

$$t=\frac{(\bar{X}_1-\bar{X}_2)-d}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

والتـــي تســـمى بـــ إحصائيـــة اختبـــار ويلتـــش Welch's Test Statistic، وهذه الإحصائيـــة كمـــا نلاحظ لها مقام مختلـــف عن إحصائية اختبـــار t الاعتيادي، وهي بشـــكلها الحالي تســـبب تضخم للخطـــأ الإحصائي من النـــوع الأول (أي تعطي نتائج غيـــر دقيقة) لذا تم تعديل درجات الحرية وفقاً لأســـلوب ســـاترثويت Satterthwaite كأحد الحلـــول لتقريب توزيع هذه الإحصائية بحيث تكون مقاربة إلى حد كبير لتوزيع t المشـــار إليه في الحالة الأولى وعليه تكون درجات الحرية وفقاً لتقريب ســـاترثويت Satterthwaite's approximation هي:

$$df=\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1}+\frac{\left(s_2^2/n_2\right)^2}{n_2-1}}$$

والاختيار بين تلك القيمتين وكما ســـبق الإشـــارة إليه يعتمد على ما إذا كان تباينا المجتمعين متســـاويين أم لا. فقد يكون معلوماً لدى الباحث مســـبقاً (مثلاً من خلال دراســـات سابقة أو من خلال الحصر الشامل) أن تبايني مجتمعي الدراسة متساويان، ومـــن ثم يتم اســـتخدام إحصائية الاختبـــار الأولى. أما إذا لم يكـــن لدى الباحث أي معلومة أو دراية مســـبقة عما إذا كان تباينا المجتمعين متســـاويين أم لا، فإنه في هذه

الحالة يســتخدم أحد أســاليب الاختبار الإحصائي لتساوي تباينات مجتمعين أو أكثر لاختبار ما إذا كان تباينا المجتمعين متســاويين أم لا من خلال بيانات عينتي الدراسة. أي اختبار الفرضية التالية:

H_0: تباينا المتغير التابع في مجتمعي الدراسة متساويان، (أي $\sigma^2_1 = \sigma^2_2$).

H_1: تباينا المتغير التابع في مجتمعي الدراسة غير متساويين، ($\sigma^2_1 \neq \sigma^2_2$).

فإذا تم «قبول» فرضية العدم فإنه يتم اســتخدام نتائج إحصائية الاختبار الأولى، أما إذا تم رفض فرضية العدم فإنه يتم استخدام نتائج إحصائية الاختبار الأخرى.

تمرين تطبيقي (٦-٤):

في دراســة للمقارنة بين متوســطي رواتب موظفي القطاعــين الحكومي والأهلي أخذت عينة عشــوائية مكونة من (٢٠) موظف لكل قطاع لهم نفس المؤهلات والخبرة العملية تقريباً ورصدت رواتبهم الشهرية بالريال السعودي كما يوضح الجدول أناه.

جدول رقم (٦-٧)
الرواتب الشهرية لموظفي القطاعين الحكومي والأهلي بالريال السعودي

القطاع الحكومي	القطاع الأهلي
٦٦٤٤	٧٧٦٦
٦١٤٢	٦٩٢٧
٦٢٩٦	٧٥٤٠
٦١٥٧	٧٤٨١
٦٧٦١	٧٣٤٥
٦٤٠٨	٧٦٨٠
٦٩٦٣	٧٤٦٠
٦٢٣٩	٧٥٧٦
٦٧٧٤	٧٤٦٠
٦٢٦٧	٧٥٤٦

تابع - جدول رقم (٦-٧).

القطاع الحكومي	القطاع الأهلي
٦٧٥٦	٧٣٨٢
٦٧٦٥	٧٤٣١
٦٧١٦	٧٤٧٤
٦٩٦٦	٧٧٠٥
٧١١٣	٧٥٧٣
٦٩٦٨	٧٣٨٨
٦٦٦٧	٧١٨٨
٦٦٥٦	٧٥٨٥
٦٢١٩	٧٥٤٩
٦٧٦٦	٧٤٨٩

المطلوب: هل هناك اختلاف معنوي بين متوسطي الرواتب في القطاعين الحكومي والأهلي؟ اختر٠٥, ٠ = α.

تمرين تطبيقي (٦-٥):

في دراسة لتأثير مادة الكافيين على عملية التمثيل الغذائي في العضلات تم اختيار ١٨ مشــاركاً من الذكور الذين خضعوا لاختبارات تمارين الذراع. تم توزيع المشــاركين عشــوائياً وبالتســاوي على مجموعتين حيث تم إعطاء المشاركين في المجموعة الأولى كبسولة تحتوي على مادة الكافيين النقي ساعة واحدة قبل الاختبار، في حين تم إعطاء المشاركين في المجموعة الثانية كبسولة الدواء الوهمي Placebo. خلال كل تمرين تم قياس معدل التبادل التنفسي (RER) ويعرف RER على أنه نسبة ثاني أكسيد الكربون CO_2 المنتجة إلى الأوكسجين O_2 المستهلكة، ومؤشراً على ما إذا كان يتم الحصول على الطاقة من الكربوهيدرات أو الدهون.

وكانت النتائج كما يلي:

جدول رقم (٦-٨)
معدل التبادل التنفسي (RER%) للمشاركين

كبسولات تحوي الكافيين	٩٦	٩٩	٩٤	٨٩	٩٦	٩٣	٨٨	١٠٥	٨٨
كبسولات وهمية	١٠٥	١١٩	١٠٠	٩٧	٩٦	١٠١	٩٤	٩٥	٩٨

المطلوب: عند مستوى معنوية $\alpha = ٠,٠٥$، هل هناك اختلاف معنوي بين متوسطي معدل التبادل التنفسي للمشاركين في المجموعتين؟

رابعاً - تحليل التباين في اتجاه واحد لمقارنة متوسطات (٣) عينات فأكثر One Way ANOVA:

الهدف من استخدامه:

لاختبار الفرضيات حول متوسطات (٣) مجتمعات فأكثر، وذلك من خلال اختبار فيما إذا كانت الفروق بين متوسطات (٣) عينات فأكثر تختلف اختلافاً معنوياً (حقيقياً) عن الفروقات بين متوسطات مجتمعات الدراسة، وهذا يتم عن طريق مقارنة مجموع التباينات بين المجموعات مع مجموع التباينات داخل المجموعات. وتجدر الإشارة إلى أنه يمكن استخدامه في حال عينتين فقط.

شروط استخدامه:

١- أن تكون العينات مستقلة.
٢- أن تكون العينات مختارة عشوائياً من مجتمعات الدراسة.
٣- أن يكون المتغير التابع محل الدراسة متصلاً (فئوياً أو نسبياً).
٤- أن يكون توزيع المتغير التابع في كل مجتمع من مجتمعات الدراسة طبيعياً.
٥- أن تكون تباينات المتغير التابع في مجتمعات (مجموعات) الدراسة متجانسة.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

سيتم توضيح كيفية تنفيذ تحليل التباين في اتجاه واحد وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٦-٤):

من ضمن فرضيات دراسة اجتماعية أجراها أحد الباحثين أن الرضا عن الحياة عند الإناث يختلف باختلاف الحالة الاجتماعية (إناث لم يتزوجن، إناث متزوجات لمرة واحدة فقط، إناث تزوجن لمرة ثانية، مطلقات). ولاختبار هذه الفرضية صمم الباحث استبانة خاصة لقياس الرضا. ومن ثم توزيع الاستبانة على عينة عشوائية مكونة من (٢٤) أنثى بعدد (٦) استبانات لكل مستوى من مستويات الحالة الاجتماعية. وكانت درجات الرضا عن الحياة حسب متغير الحالة الاجتماعية كالتالي:

جدول رقم (٦-٩)
الرضا عن الحياة عند الإناث حسب الحالة الاجتماعية

إناث لم يتزوجن	٤١	٥٤	٧٧	٩٥	٦٤	٦٩
إناث متزوجات لمرة واحدة فقط	٦٦	٤٩	٨٢	٦٢	٧٥	٩٣
إناث تزوجن لمرة ثانية	٤٣	٢٨	٥١	٤٨	٣٠	٢٣
مطلقات	٤٧	٣٦	٢٠	٢٥	٣١	٣٨

والمطلوب هو أنه عند مستوى معنوية ٠,٠٥ = α، هل يختلف متوسط الرضا عند الإناث عن الحياة حسب الحالة الاجتماعية؟

خطوات الحل:

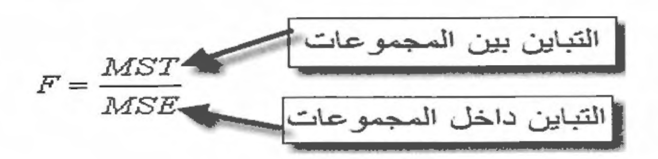
١- صياغة الفرضيات:

H_0: لا يوجد اختلاف معنوي بين متوسطات الرضا عند الإناث عن الحياة حسب الحالة الاجتماعية.

H_1: على الأقل يوجد اختلاف معنوي بين متوسطين من متوسطات الرضا عن الحياة عند الإناث حسب الحالة الاجتماعية.

٢- تحديد مستوى المعنوية: وليكن ٠,٠٥ = α.

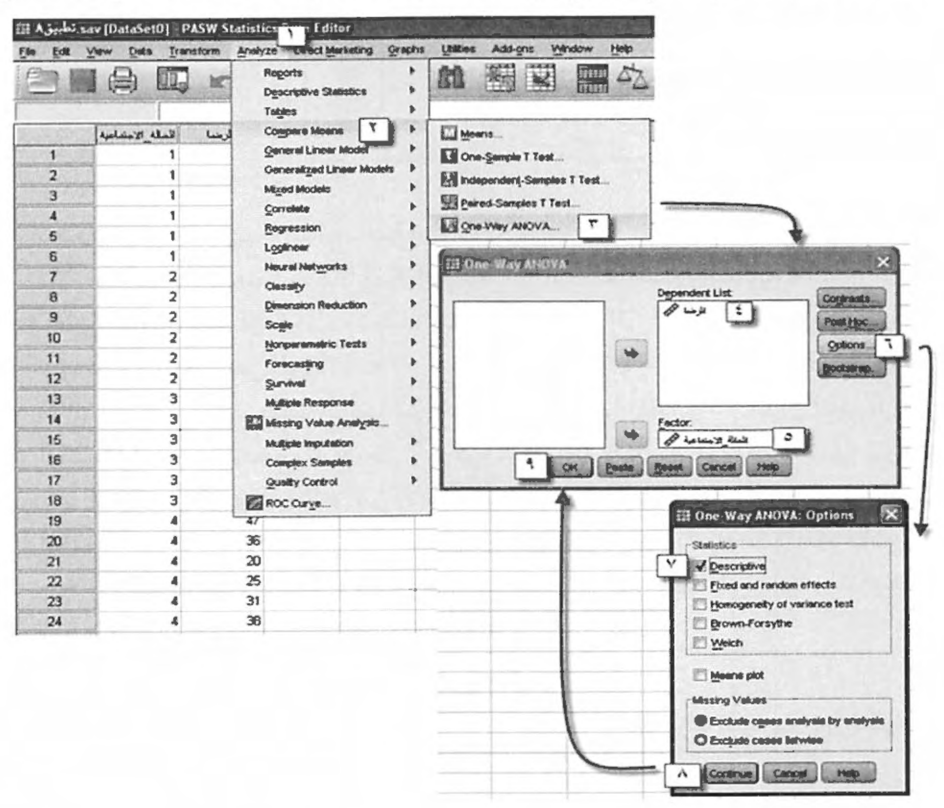
٣- تحديد الاختبار الإحصائي المناسب: الاختبار الإحصائي أو الأداة الإحصائية لاختبار فرضية العدم هو تحليل التباين في اتجاه واحد، ومن ثم تأخذ إحصائية الاختبار المستخلصة من جدول تحليل التباين الشكل التالي:



٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومـــن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حســـب الأرقام التسلسلية من ١ إلى ٩ كما هو موضح (٦-٨) لتنفيذ الاختبار.

شكل رقم (٦-٨)

خطوات تنفيذ اختبار التباين في اتّجاه واحد One-Way ANOVA لثلاث عينات فأكثر



ومن هنا نحصل على مخرجات التحليل التالية في نافذة مخرجات SPSS:

شكل رقم (٦-٩)

مخرجات تنفيذ اختبار التباين في اتجاه واحد One-Way ANOVA لثلاث عينات فأكثر

Oneway

مخرجات التحليل

[DataSet0] C:\Documents and Settings\kahtaniss\Desktop\8الخبيبة1.sav\تطبيق

Descriptives

الرضا

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
لم يتزوجن	6	66.67	18.662	7.619	47.08	86.25	41	95
متزوجات لمرة واحدة فقط	6	71.17	15.562	6.353	54.84	87.50	49	93
تزوجن لمرة ثانية	6	37.17	11.652	4.757	24.94	49.39	23	51
مطلقات	6	32.83	9.663	3.945	22.69	42.97	20	47
Total	24	51.96	21.985	4.488	42.67	61.24	20	95

p-value

ANOVA

الرضا

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7019.125	3	2339.708	11.419	0.0001
Within Groups	4097.833	20	204.892		
Total	11116.958	23			

١- اتخاذ القرار وتفسير النتائج: بما أن α (٠,٠٥) > (٠,٠٠٠١) P-value فإن الباحث سيرفض فرضية العدم. وهذا يقود إلى استنتاج أنه عند مستوى معنوية ٠,٠٥ = α فإن تحليل التباين يؤكد أن الحالة الاجتماعية للإناث تؤدي إلى الاختلاف بشكل معنوي في متوسطات الرضا عن الحياة (F (٣،٢٠) = ١١,٤٢ ، p-value < ٠,٠٠٠١). الجدول (٦-١٠) يوضح متوسطات الرضا لكل حالة وانحرافاتها المعيارية. كما أن الجدول في الشكل (٦-١٠) يوضح نتائج اختبار توكي للمقارنات البعدية للمقارنة الثنائية للرضا عن الحياة بين كل مجموعتين من الإناث في العينة حسب الحالة الاجتماعية، حيث اتضح أن هناك اختلافاً معنوياً بين المطلقات وغير المتزوجات، المطلقات والمتزوجات لمرة واحدة فقط، المتزوجات لمرة ثانية وغير المتزوجات، والمتزوجات لمرة ثانية والمتزوجات لمرة واحدة فقط، وذلك عند مستوى معنوية ٠,٠٥ = α.

جدول رقم (٦-١٠)

متوسطات الرضا عن الحياة عند الإناث حسب الحالة الاجتماعية وانحرافاتها المعيارية (بين الأقواس)

الحالة الاجتماعية			
مطلقات	تزوجن لمرة ثانية	متزوجات لمرة واحدة فقط	لم يتزوجن
٣٢,٨٣ (٩,٦٦)	٣٧,١٧ (١١,٦٥)	٧١,١٧ (١٥,٥٦)	٦٦,٦٧ (١٨,٦٧)

شكل رقم (٦-١٠)

اختبار توكي Tukey HSD للمقارنة الثنائية للرضا عند الإناث حسب حالتهن الاجتماعية

Post Hoc Tests

اختبار توكي

Multiple Comparisons

القيمة الاحتمالية المحسوبة
p-value
لمقارنة كل مجموعتين

الرضا
Tukey HSD

الحالة_الاجتماعية (I)	الحالة_الاجتماعية (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
لم يتزوجن	متزوجات لمرة واحدة فقط	-4.500-	8.264	.947	-27.63-	18.63
	تزوجن لمرة ثانية	29.500*	8.264	.010	6.37	52.63
	مطلقات	33.833*	8.264	.003	10.70	56.96
متزوجات لمرة واحدة فقط	لم يتزوجن	4.500	8.264	.947	-18.63-	27.63
	تزوجن لمرة ثانية	34.000*	8.264	.003	10.87	57.13
	مطلقات	38.333*	8.264	.001	15.20	61.46
تزوجن لمرة ثانية	لم يتزوجن	-29.500-*	8.264	.010	-52.63-	-6.37-
	متزوجات لمرة واحدة فقط	-34.000-*	8.264	.003	-57.13-	-10.87-
	مطلقات	4.333	8.264	.952	-18.80-	27.46
مطلقات	لم يتزوجن	-33.833-*	8.264	.003	-56.96-	-10.70-
	متزوجات لمرة واحدة فقط	-38.333-*	8.264	.001	-61.46-	-15.20-
	تزوجن لمرة ثانية	-4.333-	8.264	.952	-27.46-	18.80

*. The mean difference is significant at the 0.05 level.

اختبارات المقارنات في تحليل التباين Comparison Tests:

في حالة رفض فرضية العدم والوصول إلى القرار بأن هناك اختلافاً معنوياً إحصائياً في متوسطات المتغير التابع بين مجموعتين على الأقل، فإنه من المفيد تحديد أي من المجموعات مختلفة عن بعضها معنوياً في المتغير التابع. ويوجد طريقتان مختلفتان للمقارنة بين متوسطات المجموعات، ويعتمد ذلك على مدى معرفة الباحث بطبيعة العلاقة بين المجموعات. وتلك الطريقتان هما:

أ - المقارنات المخطط لها مسبقاً Preplanned Contrasts، وفيها يعد الباحث سلفاً المقارنات التي يرغب في اختبارها ولا يتم عادة إجراء جميع المقارنات الممكنة. وهذا يساعد على الحد من زيادة احتمال الوقوع في الخطأ من النوع الأول.

ب - اختبارات المقارنات البعدية post-hoc comparisons tests لاختبار الاختلافات المعنوية بين كل مجموعتين من المتوسطات. يتم عادة استخدام هذه الطريقة إذا

لم يكن لدى الباحث خلفية مسبقة عن أي المجموعات يريد مقارنتها، وإنما يعتمد على الأسلوب الاستكشافي للبحث عن المجموعات التي يوجد بينها فروقات معنوية. ومن تلك الاختبارات البعدية بونفيروني Bonferroni، فيشر لأقل فرق معنوي Fisher's LSD، ستيودنت - نيومان - كويلز Student-Newman-keuls (S-N-K) توكي Tukey، شافيه Shaffe، دنكن Duncan، دونيت Dunnett (يستخدم في حال وجود مجموعة ضابطة Control بين المجموعات) وغيرها من الاختبارات.

ج - وتجدر الإشارة إلى أن برنامج SPSS يحتوي كل تلك ١٨ اختباراً شاملاً تلك الأنواع من الاختبارات التي تم ذكرها في (ب) وغيرها من الأنواع الأخرى. والشكل (٦-١١) يوضح تلك الاختبارات مصنفة حسب تحقق شرط تجانس التباين بين مجموعات الدراسة.

شكل رقم (٦-١١)

اختبارات المقارنات البعدية المتعددة المتاحة في برنامج SPSS

والسؤال البديهي الذي يتبادر إلى ذهن القارئ هو أي تلك الاختبارات البعدية يوصى باستخدامه؟ إن اختيار الاختبار الأنسب يعتمد على عدة نقاط وهي:

- هدف المقارنة، بمعنى هل يرغب الباحث مثلاً في مقارنة المجموعة الضابطة مع أي مجموعة أخرى، أم يرغب في مقارنات ثنائية بين كل مجموعتين، أم غير ذلك.
- درجة ضبط احتمالية الوقوع في الخطأ من النوع الأول.

- درجة ضبط احتمالية الوقوع في الخطأ من النوع الثاني (قوة الاختبار المطلوبة).
- هــل نتائج اختبار المقارنة موثوق بها فيما لــو كان هناك مخالفة لأحد الافتراضات اللازمة لاختبار تحليل التباين ANOVA – وتحديداً هل تباينات مجموعات الدراسة متجانسة أم لا؟

وقــد ناقــش كل مــن دوقــلاس مونتقمــري (2012) Montgomery وآندي فيلد (2005) Field وهوك (2012) Huck وغيرهم تلك الاختبارات وخصائصها والحالات التي يمكن يفضل استخدام تلك الاختبارات فيها بطريقة موجزة مستخلصة من نتائج دراســات في هذا المجال. والاختيار من بين تلك الاختبارات يعتمد على الموازنة قدر الإمكان بين النقاط الســابق ذكرها. ولقد أشــار فيلد أن اختبــار توكي وبونفيروني يعملان على ضبط درجة الوقوع في الخطأ من النوع الأول بشكل جيد، ولكن هذا على حســاب فقدان بعض من قوتهما الإحصائية Statistical power. كما بين أن اختبار توكــي أكبر قوة من اختبــار بونفيروني فيما إذا كان عدد المقارنات كبيراً نســبياً وفي المقابــل بونفيروني أكبر قوة من توكــي فيما إذا كان عدد المقارنات أقل. وأوضح فيلد بأن اختبار توكي وبونفيروني أقوى إحصائياً (القدرة على اكتشــاف الفروق الحقيقية أو المعنوية بين متوســطات المجموعات) من اختباري دونكن وشافيه. علاوة على ذلك، فقد اقترح فيلد اســتخدام اختبار قيمز – هاويل Games-Howell، بالإضافة إلى أي اختبار آخر حســب ما يراه الباحث عندما يكون هناك شــك في عدم تجانس تباينات مجتمعات الدراسة لما يتمتع به هذا الاختبار من أداء جيد بوجه عام.

وذكر هوك (2012) Huck أن اختبــارات بونفيروني Boneferroni، توكي Tukey HSD وشــافيه Shaffe هي أكثر اختبــارات المقارنات البعدية اســتخداماً، في حين اختبــارات فيشــر Fisher's LSD، ودنكــن Duncan، ونيومان – كويلز – ســتيودنت N-K-S أقل استخداماً.

كمــا قدم قــاري إنقرســول (2010) Ingersoll مجموعة من الإرشــادات لاختيار الاختبار الأنســب للمقارنة بين متوســطات المجموعات، حيث اقترح استخدام اختبار Tukey إذا كان الباحــث يريد إجراء جميع المقارنــات الثنائية الممكنة بين المجموعات في الدراســة، وهذه الرغبة تكون عادة مرتبطــة بعدم معرفة الباحث بأي المجموعات يريد مقارنتها دون الأخرى. كما أوصى أغريستي وفنلي (2009) Agresti & Finlay باستخدام أحد الاختبارين بونفيروني أو توكي.

توصيات حول اختيار اختبار المقارنة البعدية المناسب:

وبناء على ما سبق أضع بين يدي القارئ الاقتراحات والتوصيات التالية المتعلقة باختيار اختبار المقارنات البعدية المناسب وهي كالتالي:

- إذا كانت شروط اختبار تحليل التباين متحققة يتم استخدام اختبار توكي كخيار أول Tukey HSD، أو شافيه Shaffe، أو اختبار R-E-G-W-Q كخيار ثاني.
- لزيادة مستوى الضمان من عدم الوقوع في الخطأ من النوع الأول يتم اختيار اختبار بونفيروني Boneferroni وذلك في حالة العدد القليل من المقارنات الثنائية.
- في حالة عدم تحقق شرط تجانس تباينات المجموعات يتم استخدام اختبار قيمز - هاويل Games-Howell.
- في حالة عدم تساوي أحجام العينات في مجموعات المقارنة فإنه يوصى باستخدام اختبار هوك بيرغ Hochberg's GT2 إذا كان الاختلاف بين أحجام العينات كبير، أو استخدام اختبار غابريل Gabriel في حال كان الاختلاف بين أحجام العينات صغيراً.
- أما إذا أراد الباحث مقارنة مجموعة ضابطة Control Group بمجموعات أخرى فإنه يوصى باستخدام اختبار دونيت Dunnett الذي صمم لهذا الغرض.

إضاءات إحصائية حول الاختبارات المعلمية للفروق بين المتوسطات:

١- يمكن للباحث دراسة الفروق بين متوسطات متغير تابع في ظل وجود عاملين أو أكثر (متغيرين مستقلين أو أكثر) في آن واحد، وذلك باستخدام تحليل التباين في اتجاهين أو أكثر N-Way ANOVA، ومثال على ذلك اختبار تأثير المؤهل التعليمي ومستوى الخبرة على الرضا الوظيفي لدى الموظفين.

٢- لا ينصح باستخدام اختبار t للمقارنة بين أكثر من مجموعتين بحيث تتم مقارنة كل مجموعتين على حدة، لأن ذلك يؤدي إلى زيادة احتمال تضخم الخطأ من النوع الأول. أي زيادة احتمال الوصول إلى قرار خاطئ بأنه يوجد اختلاف معنوي بين متوسطي مجموعتين مستقلتين مع العلم أنها في الواقع غير ذلك.

تمرين تطبيقي (٦-٦):

الجدول التالي يوضح درجات الرضا الوظيفي لعينة عشوائية من إحدى الجهات الحكومية حسب مستوى خبراتهم، وذلك ضمن دراسة ميدانية عن الرضا الوظيفي. وكان من ضمن تساؤلات الباحث هو الكشف عن العلاقة بين درجة الرضا الوظيفي ومستويات الخبرة (أقل من ٥ سنوات، ٥-١٠سنوات، أكثر من ١٠ سنوات).

جدول رقم (٦-١١)
درجات الرضا الوظيفي حسب مستوى الخبرة

الخبرة	درجة الرضا الوظيفي							
أقل من ٥ سنوات	١٢	١٠	١٤	٩	١٣	١٠	١٠	١١
من ٥-١٠ سنوات	١٣	١٥	١٦	١٢	١٤	١٧	١٤	١٥
أكثر من ١٠ سنوات	١٥	١٨	١٨	١٤	١٩	١٧	١٦	١٨

والمطلوب: اختبر ما إذا متوسط الرضا الوظيفي لموظفي تلك الجهة الحكومية يختلف باختلاف مستويات الخبرة الوظيفية عند مستوى معنوية ٠,٠٥ =α.

تمرين (٦-٧):

افترض باحث في دراسة أن إنتاجية الموظف تختلف حسب المؤهل العلمي (متوسطة، ثانوية، بكالوريوس فأعلى). لذا قام الباحث بتصميم استبانة تم فيها قياس الإنتاجية والجدول التالي يوضح درجات الإنتاجية لـ (٢٤) موظفاً حسب المؤهل العلمي. هل تؤيد هذه البيانات افتراض الباحث عند مستوى معنوية ٠,٠٥ =α؟

جدول رقم (٦-١٢)
إنتاجية الموظف حسب المؤهل العلمي

متوسطة	٦٠	٥٣	٥٤	٥١	٦٠	٦٥	٧٠	٥٥
ثانوية	٧٠	٦٦	٦٥	٧٠	٨٠	٨١	٨٥	٨٠
بكالوريوس فأعلى	٨٠	٧٠	٦٠	٧٥	٨٠	٨٨	٩٠	٩١

الفصل السابع
الاختبارات اللامعلمية للفروق بين المتوسطات
NonParametric Tests for Means Differences

مقدمة:

في حالة عدم ملاءمة استخدام الاختبار المعلمي للفروق بين المتوسطات لسبب أو أكثر من الأسباب التي تم ذكرها في الفصل الرابع مثل الالتواء الشديد في توزيع البيانات وعدم تجانس تباينات المتغير التابع بين مجتمعات الدراسة وغيرذلك، يمكن استخدام البديل اللامعلمي لتنفيذ الاختبار؛ لأنه في هذه الحالة أقدر على اكتشاف الفروق الحقيقية بين المتوسطات إن وجدت. وفي هذا الفصل سيتم تقديم أكثر الاختبارات الإحصائية اللامعلمية البديلة والمناظرة للاختبارات المعلمية التي تم تقديمها في الفصل السابق لاختبار الفرضيات حول المتوسطات وشرحها بطريقة عملية تفاعلية باستخدام برنامج SPSS (القحطاني، ١٤٣٢).

أولاً - اختبار ويلكوكسن للرتب لعينة واحدة One-Sample Wilcoxon Signed-Ranks Test:

الهدف من استخدامه:

لاختبار الفرضية حول وسيط مجتمع واحد، وذلك من خلال اختبار ما إذا كان وسيط العينة يختلف اختلافاً معنوياً (حقيقياً) عن القيمة المفترضة لوسيط المجتمع.

متى يستخدم:

يستخدم كبديل لاختبار t (المعلمي) لمتوسط عينة واحدة عندما يكون توزيع المتغير التابع غير طبيعي وحجم العينة صغير أو عندما تكون البيانات رتبية وحجم العينة صغير.

شروط استخدامه:

١- أن تكون المتغير التابع (متغير الاستجابة) في الدراسة تمثل عينة عشوائية من مجتمع الدراسة ومستقلة بعضهاعن بعض.

٢- أن يكون المتغير التابع محل الدراسة رتبياً على الأقل.

٣- أن يكــون توزيــع المتغير التابع في مجتمع الدراســة متماثــلاً، لذلك يمكن اختبار المتوسط الحسابي بدلاً من الوسيط.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي ســيتم توضيح كيفية تنفيذ اختبار ويلكوكسن للرتب لوسيط عينة واحدة وتفسير النتائج من خلال التطبيق العملي التفاعلي التالي:

تطبيق عملي تفاعلي (٧-١):

في إحدى الدراســات المتعلقة بوجهات نظر الموظفين حــول جوانب متعددة تتعلق بالوظيفة، تم توزيع اســتبانة مكونة من عدد من الأســئلة ذات الصلة بالموضوع محل البحث. والبيانات في الجدول (٧-١) تمثل إجابة (٣٠) موظفاً تم اختيارهم عشوائياً من مجتمع الدراســة عن ســؤالي الاهتمام بالأمن الوظيفي والاهتمام بالمقابل المادي الذي يحصلون عليه. وقد تم اســتخدام مقيــاس متدرج من ١ (تعني لا يوجد اهتمام مطلقاً) إلى ١٠ (تعني اهتماماً بالغاً جداً) لتمثيل درجة الاهتمام للموظف.

جدول رقم (٧-١)
درجة الاهتمام بالأمن الوظيفي والمقابل المادي للموظفين

الموظف	درجة الاهتمام بالأمن الوظيفي	درجة الاهتمام بالمقابل المادي
١	٣	٤
٢	١	٩
٣	٦	٨
٤	٧	٥
٥	٦	٤
٦	٥	٦
٧	٣	٦
٨	٤	٤
٩	٨	٧

تابع - جدول رقم (٧-١).

الموظف	درجة الاهتمام بالأمن الوظيفي	درجة الاهتمام بالمقابل المادي
١٠	٣	٥
١١	٤	٥
١٢	٦	٨
١٣	٦	٧
١٤	٣	٦
١٥	٧	٧
١٦	٥	٨
١٧	٤	٥
١٨	٤	٣
١٩	٦	٥
٢٠	٦	٧
٢١	١	٥
٢٢	٧	٣
٢٣	٤	٥
٢٤	٥	٦
٢٥	٤	٤
٢٦	٢	٥
٢٧	٦	٦
٢٨	٤	٥
٢٩	٣	٦
٣٠	٢	٦

هل تدل البيانات على ارتفاع أو انخفاض في درجة الاهتمام بالأمن الوظيفي لدى الموظفين؟ بمعنى هل تقل أو تزيد درجة الاهتمام بالأمن الوظيفي بشكل معنوي عن وسيطه؟

خطوات الحل:

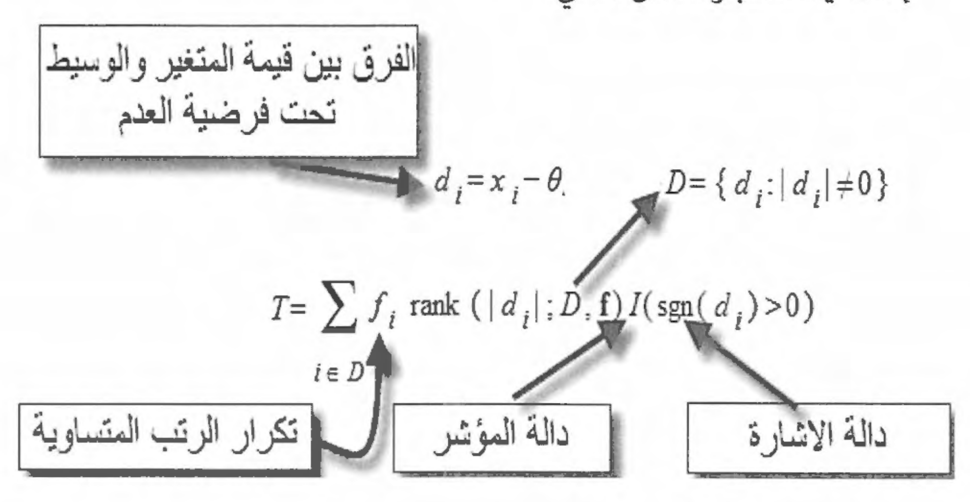
١- صياغة الفرضيات:

H_0: وسيط درجة الاهتمام بالأمن الوظيفي = ٥,٥ (أي لا يوجد اختلاف - ارتفاع أو انخفاض عن الوسيط - معنوي في درجة الاهتمام بالأمن الوظيفي).

H_1: وسـيط درجة الاهتمام بالأمن الوظيفي ≠ ٥,٥ (يوجد اختلاف - ارتفاع أو انخفاض عن الوسيط - معنوي في درجة الاهتمام بالأمن الوظيفي).

٢- تحديد مستوى المعنوية: وليكن $\alpha =$ ٠,٠٥.

٣- تحديد الاختبار الإحصائي المناسب: في هذه الحالة سيتم اختيار اختبار ويلكوكسن للرتب لعينة واحدة Wilcoxon Signed-Rank Test One-Sample ومن ثم تأخذ إحصائية الاختبار الشكل التالي:

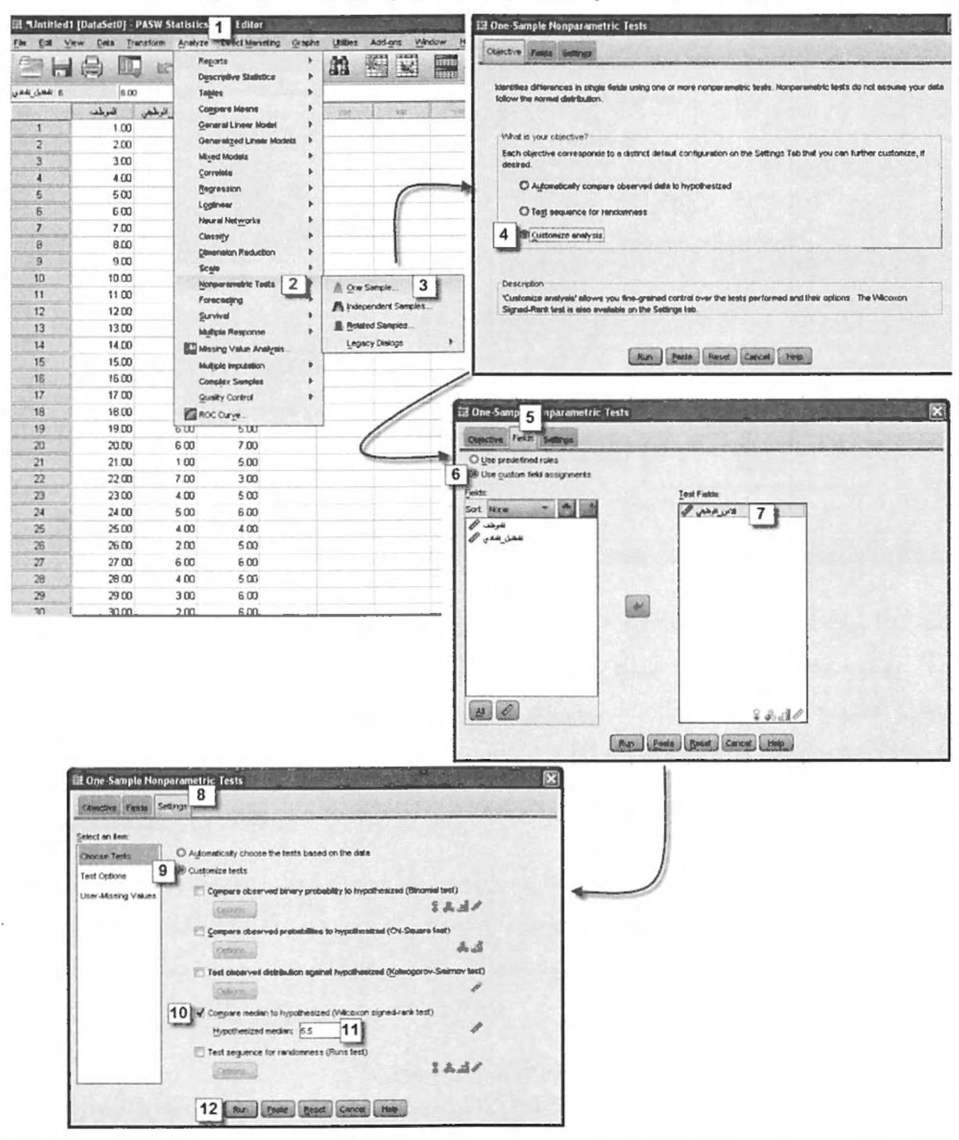


٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومـن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسـب الأرقام التسلسلية من ١ إلى ١٢ كما هو موضح بالشكل (٧-١) لتنفيذ الاختبار.

ملاحظـة: يتـم تحويل إحصائيـة الاختبار T إلى الصيغـة المعيارية Z وذلك بقسـمة الفرق بين T ومتوسطها على الانحراف المعياري لـ T ومن ثم استخدامها في حسـاب القيمة الاحتمالية التي يتم مقارنتها بمستوى المعنوية α لاتخاذ القرار الإحصائي حيال فرضية العدم بالقبول أو عدم القبول.

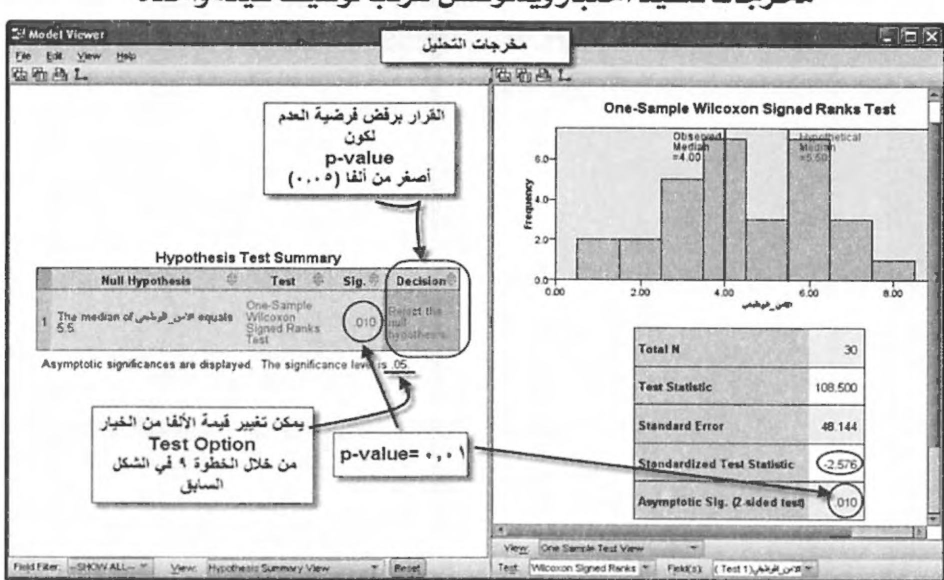
شكل رقم (٧-١)

خطوات تنفيذ اختبار ويلكوكسن للرتب لوسيط عينة واحدة



وبعـــد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output، نقوم بالنقر عليه مرتين متتاليتين للحصول على التالي:

شكل رقم (٧-٢)
مخرجات تنفيذ اختبار ويلكوكسن للرتب لوسيط عينة واحدة



٥- اتخاذ القرار وتفسير النتائج: بما أن (٠,٠٥) α > (٠,٠١) P-value فإن الباحث سيقوم برفض فرضية العدم ومن ثم قبول الفرضية البديلة. وهذا يقود إلى استنتاج أنه عند مستوى معنوية ٠,٠٥ = α فإن بيانات العينة تبين أن وسيط درجة الاهتمام بالأمن الوظيفي (٤,٠٠) منخفض، (Z = -2.58, p-value = 0.01).

تمرين تطبيقي (٧-١):

في دراسة عن مستوى رضا المراجعين عن خدمات الرعاية الصحية الأولية التي يقدمها مركز العليا والسليمانية في مدينة الرياض، تم اختيار عينة عشوائية مكونة من (٦٦) من المراجعين لهذا المركز وتم سؤالهم عن مستوى رضاهم عن خدمات المركز وكانت النتائج كالتالي (فهمي، ٢٠٠٥، ٣٦٤):

جدول رقم (٧-٢)
مستوى رضا المراجعين عن خدمات الرعاية الصحية الأولية التي يقدمها مركز العليا والسليمانية بالرياض

درجة الرضا	غير راضٍ تماماً (١)	غير راضٍ (٢)	متوسط الرضا (٣)	راضٍ (٤)	راضٍ تماماً (٥)
عدد المراجعين	١٠	١٣	١٣	٢٠	١٠

المطلوب: هل تدل هذه البيانات على ارتفاع أو انخفاض درجة رضا المراجعين لهذا المركز عن خدمات الرعاية الأولية التي يقدمها؟ (اختر $\alpha = ٠,٠٥$).

تمرين تطبيقي (٧-٢):

صـــدر تقرير عن مكتـــب الولايات المتحدة الامريكية لإحصـــاءات العمل يفيد بأن متوســـط عدد الوظائف التي يمارســـها الفرد في الفئة العمرية ما بين ١٨ و٣٤ عاماً بلغت تقريبا ٩,٢ وظيفة. وللتحقق من صحة ذلك، قام باحث باختيار ٨ أفراد يعملون من الفئة العمرية بين ١٨ و٣٤ عاماً، وسألهم عن عدد الأماكن المختلفة التي عملوا بها، وكانت النتائج كما هي موضحة بالجدول (٧-٣) أدناه:

جدول رقم (٧-٣)
عدد الوظائف لعينة عشوائية مكونة ٨ أفراد تتراوح أعمارهم بين ١٨ و٣٤ عاماً

العامل	١	٢	٣	٤	٥	٦	٧	٨
عدد الوظائف	٨	١٢	١٥	٦	١	٩	١٣	٢

المصدر: (http://www.chegg.com/homework-help/questions-and-answers/us-bureau-labor-andstatistics-reported-person-ages-of18-34-average-92-jobs-see-thisaverage-q1380158)

والمطلوب هو: بناء على بيانات العينة وعند مستوى معنوية $\alpha = ٠,٠٥$، هل يختلف متوسط عدد الوظائف عن ٩,٢ وظيفة للفرد بين الفئة العمرية بين ١٨ و٣٤؟

ثانياً - اختبار ويلكوكسن للرتب لوسيطي عينتين مرتبطتين Wilcoxon Signed-rank Test Related-Sample:

الهدف من استخدامه:

لاختبـــار الفرضية حول وســـيطين مجتمعين مرتبطين. أي اختبـــار فيما إذا كان وسيطا مجتمعي الدراسة متساويين أم مختلفين معنوياً.

متى يستخدم:

يستخدم اختبار ويلكوكســـن اللامعلمي بديلاً لاختبار t (المعلمي) لمتوسطي عينتين مرتبطتين، وذلك عند مخالفة اختبار t لأحد شروط استخدامه مثل كون توزيع الفروق بين قيم المتغير التابع في مجتمعي الدراســـة غير طبيعي أو لكون بيانات المتغير التابع رتبية في ظل وجود عينة صغيرة الحجم.

شروط استخدامه:

١- أن تكون العينتان مرتبطتين.

٢- أن تكون العينتان مختارتين عشوائياً من مجتمعي الدراسة.

٣- أن يكــون المتغيــر التابــع محــل الدراســة رتبياً علــى الأقل (يعني رتبــي، فتري أو نسبي).

٤- أن يكــون توزيع الفروق بين قيم المتغير التابع في مجتمعي الدراســة متماثلاً حول وسيطها (لذلك يمكن اختبار المتوسط عوضاً عن الوسيط).

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي ســيتم توضيح كيفية تنفيذ اختبار ويلكوكســن للرتب لوســيطي عينتين مرتبطتين وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي:

تطبيق عملي تفاعلي (٧-٢):

بالعــودة إلى التطبيق العملي التفاعلي (٧-١)، هــل درجة اهتمام الموظفين بالأمن الوظيفــي مختلفــة معنوياً عــن اهتمامهم بالمقابل المــادي الذي يحصلــون عليه من الوظيفة؟

خطوات الحل:

١- صياغة الفرضيات:

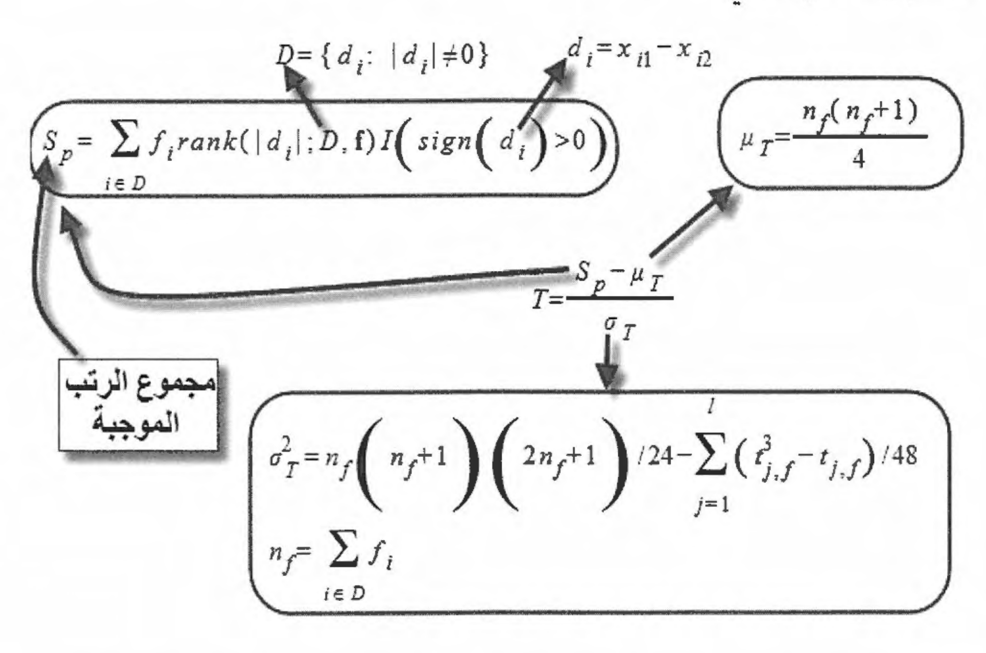
H_0: لا يوجد اختلاف معنوي بين وسيطي درجة اهتمام الموظفين بالأمن الوظيفي ودرجة اهتمامهم بالمقابل المادي الذي يحصلون عليه من الوظيفة.

H_1: يوجد اختلاف معنوي بين وســيطي درجة اهتمام الموظفين بالأمن الوظيفي ودرجة اهتمامهم بالمقابل المادي الذي يحصلون عليه من الوظيفة.

٢- تحديد مستوى المعنوية: وليكن $\alpha =$ ٠,٠٥.

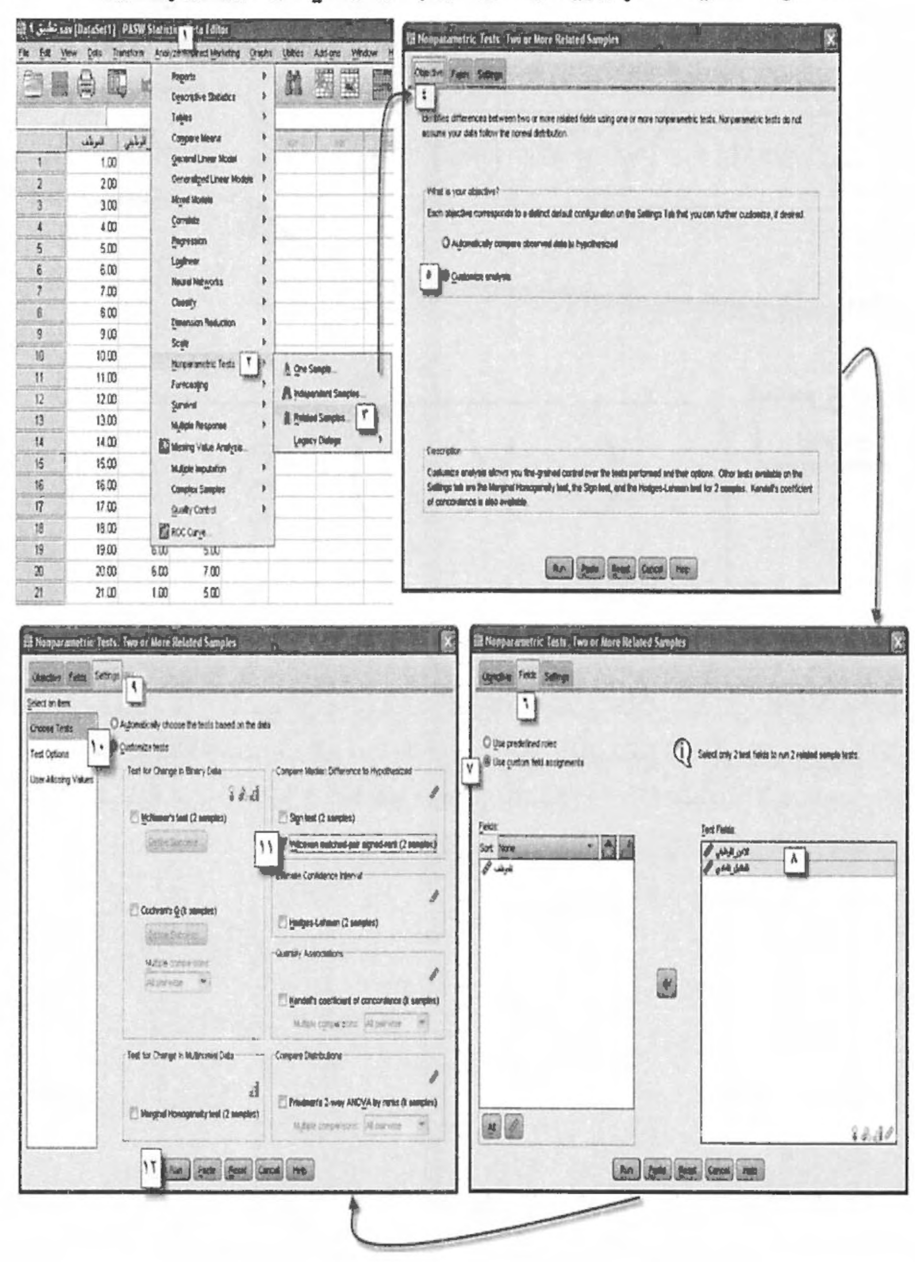
٣- تحديد الاختبار الإحصائي المناسب: الاختبار الإحصائي المناسب في هذه الحالة هو ويلكوكسن للرتب لوسيطي عينتين مرتبطتين.

(Related-Samples Wilcoxon Signed-rank Test) وبالتالـي تأخـذ إحصائية الاختبار الشكل التالي:



٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومـن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسـب الأرقام التسلسلية من ١ إلى ١٢ كما هو موضح بالشكل (٧-٣) لتنفيذ الاختبار.

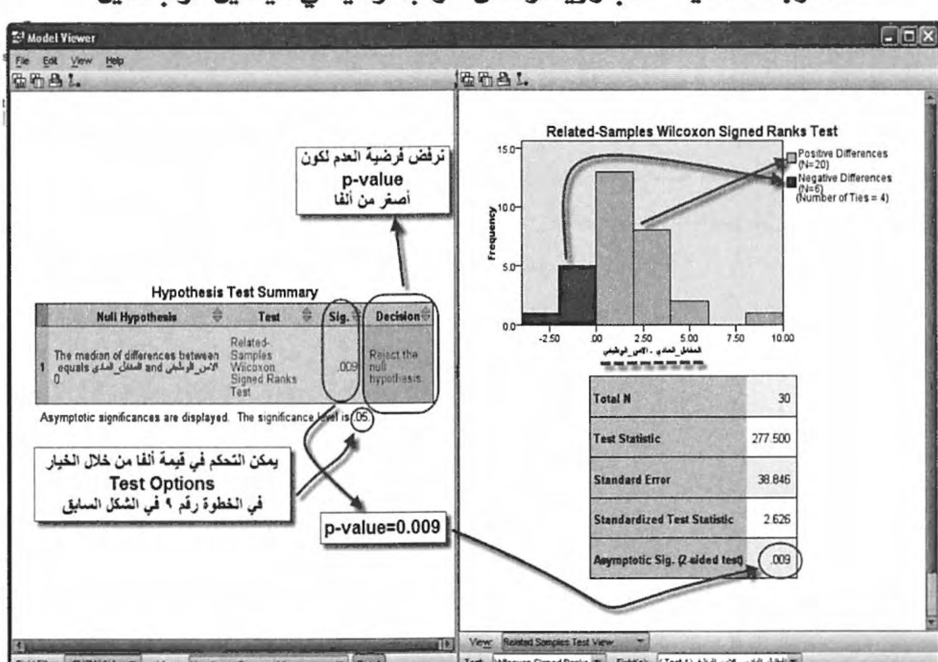
شكل رقم (٧-٣)
خطوات تنفيذ اختبار ويلكوكسن للرتب لوسيطي عينتين مرتبطتين



وبعد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output، نقوم بالنقر عليه مرتين متتاليتين للحصول على التالي:

شكل رقم (٧-٤)

مخرجات تنفيذ اختبار ويلكوكسن للرتب لوسيطي عينتين مرتبطتين



٥- اتخاذ القرار وتفسير النتائج: بما أن $(\cdot,\cdot\cdot٩) < \alpha\ (\cdot,\cdot٥)$ P-value فإن الباحث سيقوم برفض فرضية العدم ومن ثم قبول فرضية البدل. وهذا يقود إلى استنتاج أنه عند مستوى معنوية $\alpha = \cdot,\cdot٥$ فإن بيانات العينة تدل على أن وسيط درجة الاهتمام بالمقابل المادي (٥,٥) يختلف معنوياً عن وسيط درجة الاهتمام بالأمن الوظيفي (٤,٠٠) (Z = 2.63, p-value = 0.009).

تمرين تطبيقي (٧-٣):

لمعرفة مدى تأثير إشارة ضوئية جديدة في تقليص عدد الحوادث المرورية، تم الحصول على البيانات التالية التي تمثل عدد الحوادث في ١٢ تقاطعاً خطراً خلال أربعة أسابيع قبل وبعد تركيب الإشارة الضوئية:

جدول رقم (٧-٤)

عدد الحوادث في ١٢ تقاطعاً خطراً خلال أربعة أسابيع قبل وبعد تركيب الإشارة الضوئية

التقاطع	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
قبل	٢	٣	٢	١	٢	٦	٥	٤	٥	٣	٢	٤
بعد	١	٢	٠	٣	١	٣	٣	١	٢	٢	٣	٢

والسؤال هو عند مستوى دلالة $\alpha = ٠,٠٥$ هل تدل البيانات على أن هناك اختلافاً معنوياً في عدد الحوادث قبل وبعد تركيب الإشارة الضوئية (بمعنى هل وجود الإشارة الضوئية عند تقاطع الطرق يسهم في خفض عدد الحوادث)؟

ثالثاً - اختبار مان - ويتني U لوسيطي عينتين مستقلتين Mann-Whitney U Test:

الهدف من استخدامه:

لاختبار الفرضية حول وسيطي مجتمعين مستقلين. أي اختبار ما إذا كان وسيطا مجتمعي الدراسة متساويين أم مختلفين معنوياً.

متى يستخدم:

يستخدم اختبار مان – ويتني U اللامعلمي بديلاً لاختبار t المعلمي لمتوسطي عينتين مستقلتين، وذلك عند مخالفة اختبار t لأحد شروط استخدامه مثل كون توزيع المتغير التابع في مجتمعي الدراسة غير طبيعي أو عدم تجانس تباينات المجموعتين أو لكون بيانات المتغير التابع رتبية، وذلك في ظل وجود عينات صغيرة الحجم.

شروط استخدامه:

١- أن تكون العينتان مستقلتين.

٢- أن تكون العينتان مختارتين عشوائياً من مجتمعي الدراسة.

٣- أن يكون المتغير التابع محل الدراسة رتبياً على الأقل (يعني رتبي، فتري أو نسبي).

٤- أن يكون توزيع المتغير التابع في مجتمعي الدراسة متطابقين في الشكل.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سيتم توضيح كيفية تنفيذ اختبار مان – ويتني U لوسيطي عينتين مستقلتين وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي:

تطبيق عملي تفاعلي (٧-٣):

أراد مدير إحدى المنظمات أن يتعرف على ما إذا كان هناك فروق معنوية بين الموظفين والموظفات في درجة رضاهم الوظيفي أم لا، لذا قام باختيار عينتين عشوائيتين مكونتين من (١٢) من الموظفين الذكور و(١٢) من الموظفات. وطلب منهم أن يُبدوا درجة رضاهم الوظيفي على مقياس يتراوح ما بين (١) غير راضٍ إطلاقاً إلى (١٠) راضٍ تماماً، وكانت إجاباتهم كالتالي:

جدول رقم (٧-٥)
درجة الرضا الوظيفي للموظفين والموظفات

المجموعة	درجة الرضا الوظيفي											
الموظفون	٨	٧	٦	٩	١٠	٤	٥	٩	١٠	٧	٦	٨
الموظفات	٣	٤	١	٩	٨	٥	٨	٩	٣	٤	–	–

هل تدل بيانات العينة على وجود اختلاف معنوي في درجة الرضا الوظيفي بين الموظفين والموظفات عند مستوى معنوية $\alpha = ٠,٠٥$؟

خطوات الحل:

١- صياغة الفرضيات:

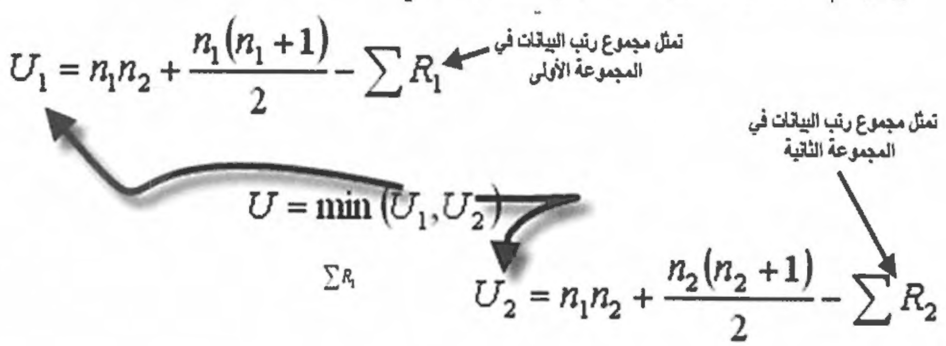
H_0: لا يوجد اختلاف معنوي بين وسيطي درجة الرضا الوظيفي للموظفين والموظفات.

H_1: يوجد اختلاف معنوي بين وسيطي درجة الرضا الوظيفي للموظفين والموظفات.

٢- تحديد مستوى المعنوية: وليكن $\alpha = ٠,٠٥$.

٣- تحديد الاختبار الإحصائي المناسب: الاختبار الإحصائي في هذه الحالة هو مان – ويتني U لوسيطي عينتين مستقلتين (Mann-Whitney U Test).

ومن ثم تأخذ إحصائية الاختبار الشكل التالي:



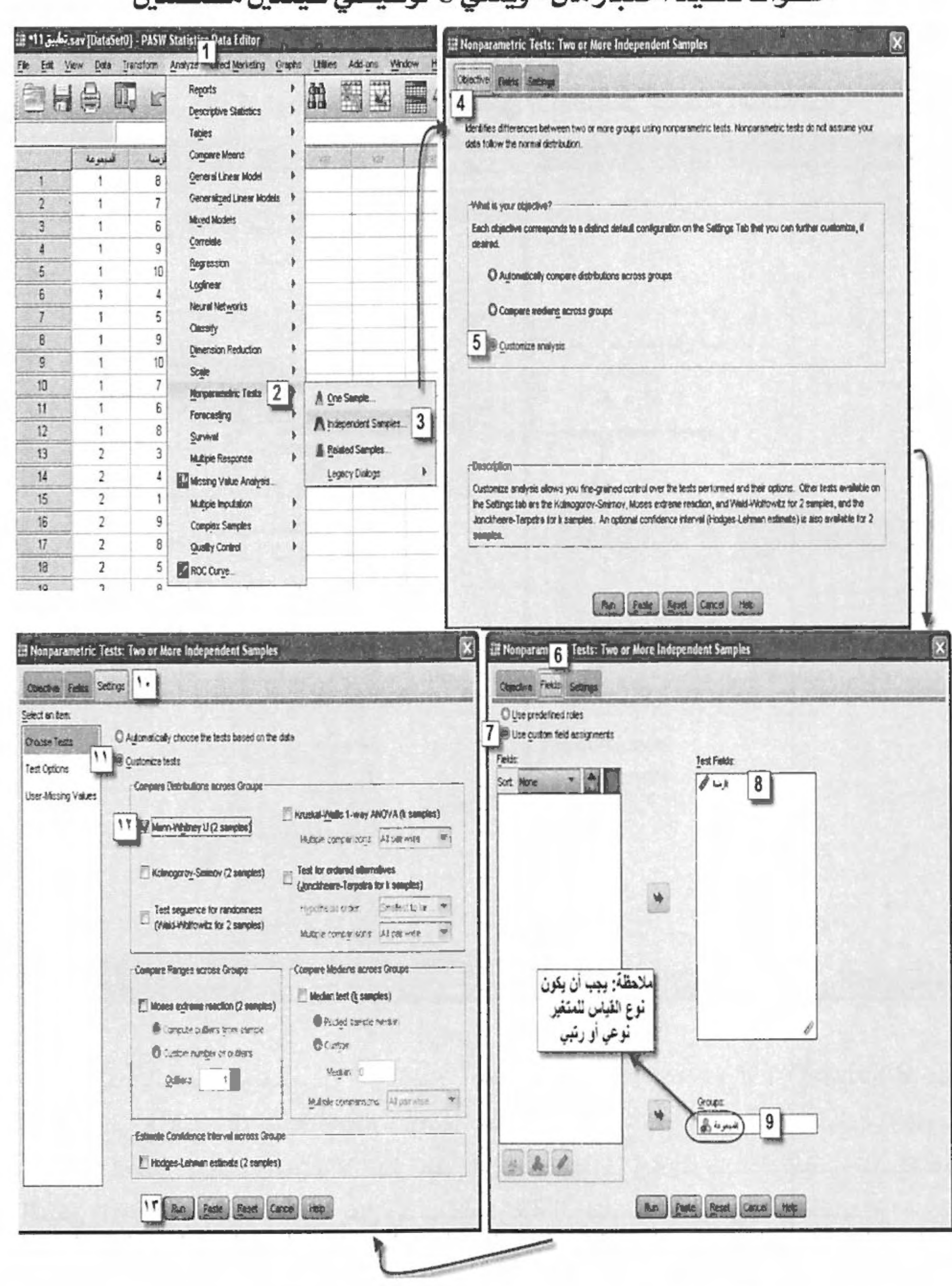
وتجدر الإشارة إلى ما يلي:

- يوجد صياغة مختلفة لإحصائية اختبار مان - ويتني يمكنك الاطلاع عليها من خلال دليل SPSS في قائمة Help.
- يتم تحوير إحصائية الاختبار U في حال وجود رتب متساوية وفي حال استخدام توزيعات احتمالية أخرى كالتوزيع الطبيعي المعياري Z كتقريب لتوزيع U.
- هناك اختبار آخر يمكن أيضاً استخدامه بدلاً من اختبار مان - ويتني وهو اختبار الوسيط Median Test فهو يعتبر شائع الاستخدام أيضاً.

٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ١٣ كما هو موضح بالشكل (٧-٥) لتنفيذ الاختبار.

شكل رقم (٧-٥)

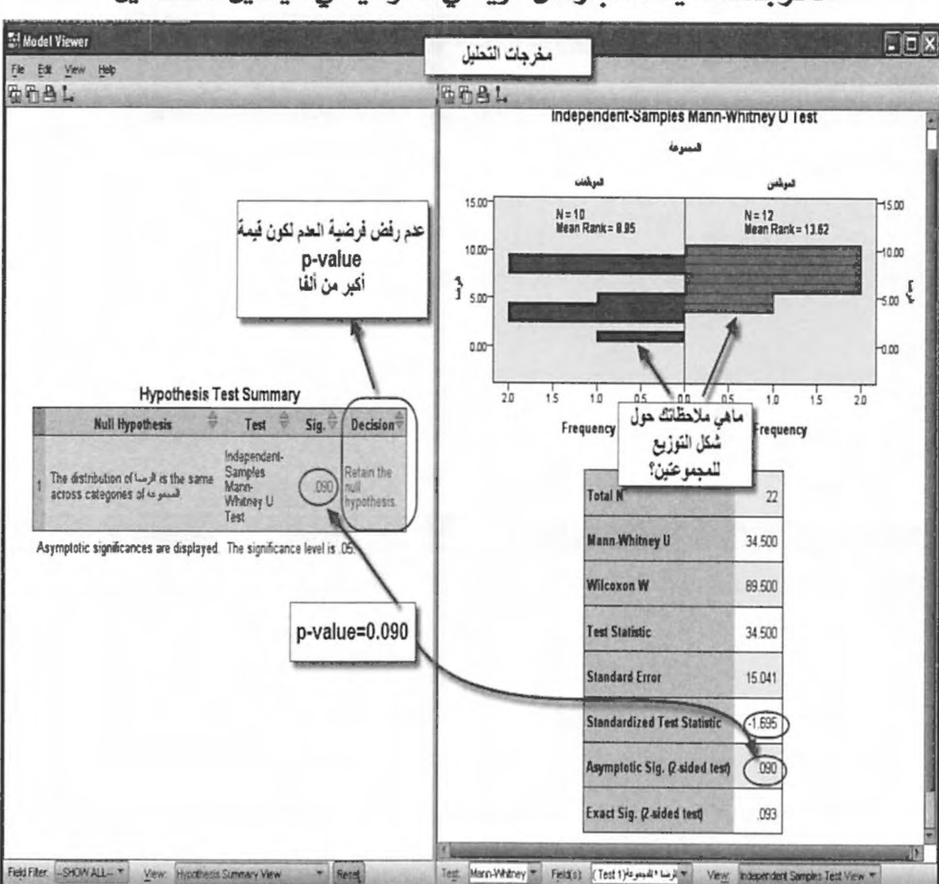
خطوات تنفيذ اختبار مان - ويتني U لوسيطي عينتين مستقلتين



وبعـــد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output، نقوم بالنقر عليه مرتين متتاليتين للحصول على التالي:

شكل رقم (٧-٦)

مخرجات تنفيذ اختبار مان - ويتني U لوسيطي عينتين مستقلتين



اتخـــاذ القـــرار وتفســـير النتائـــج: بمـــا أن P-value (٠,٠٩) > α (٠,٠٥) فإن الباحث لن يرفض فرضية العدم. وهذا يقود إلى اســـتنتاج أنه عند مســـتوى معنوية α = ٠,٠٥ فإن بيانات العينة لا تدل على أن وســـيط درجة الرضا الوظيفي للموظفين الذكور (٧,٥) يختلف معنوياً عن وســـيط درجة الرضـــا الوظيفي للموظفات (٤,٥٠) .(Z = -1.695, p-value = 0.09)

تمرين تطبيقي (٧-٤):

في دراســة اســتطلاعية لمقارنة تأثير طريقتين مختلفين مــن الإعلانات التجارية (X وY) على شــراء المنتجــات، قام أحد الباحثين في مجال التســويق بمقارنة تأثير تلــك الطريقتين من الإعلانات التجارية على شــراء منتج جديد من أجهزة الحاســب الآلي الشــخصية. حيث تم اختيار عينتين عشوائيتين مستقلتين من العملاء من أحد المحــلات التجارية الكبرى المتخصصة في بيع الأجهزة الإلكترونية، وتم تعريض إحدى المجموعتــين للنوع الأول من الدعاية والأخرى للنــوع الآخر من الدعاية. بعد ذلك تم طلبهم في اختيار رقم واحد فقط من ١ (لن يشــتري المنتج) إلى ١٠ سيشــتري المنتج بالتأكيد) للتعبير عن درجة احتمالية شراء المنتج الجديد. والجدول (٧-٦) التالي يبين نتيجة الاستطلاع.

جدول رقم (٧-٦)
تأثير طريقة الإعلانات التجارية على شراء منتج جديد

الطريقة	درجة احتمالية شراء المنتج الجديد من الحاسبات الشخصية					
X	٣	٤	٢	٦	٢	٥
Y	٩	٧	٥	١٠	٦	٨

هــل هناك اختلاف معنوي بين الطريقتين المختلفتين مــن الإعلانات التجارية في درجة احتمالية الشراء؟ اختر مستوى معنوية محدد مثل $\alpha =$ ٠,٠٥؟

تمرين تطبيقي (٧-٥):

البيانات التالية تمثل الأعمار التي عندها تم تشــخيص مجموعتين عشوائيتين من البالغين الصغار من الذكور والإناث بمرض السكري من النوع الثاني.

جدول رقم (٧-٧)
توزيع الأشخاص المصابين بمرض السكري من النوع الثاني حسب الجنس والعمر

الجنس	العمر				
ذكر	١٩	٢٢	١٦	٢٩	٢٤
أنثى	٢٠	١١	١٧	١٢	١٣

والمطلوب: هل متوسط (وسيط) الأعمار عند التشخيص بالمرض يختلف معنوياً باختلاف الجنس؟ حدد مستوى للمعنوية وليكن $\alpha = ٠,٠٥$؟

رابعاً - اختبار كروسكال - واليس لأوساط ٣ عينات مستقلة فأكثر Kruskal-Wallis Test:

الهدف من استخدامه:

لاختبار الفرضية حول أوساط (جمع وسيط) ثلاث مجتمعات مستقلة أو أكثر. أي اختبار فيما إذا كانت أوساط مجتمعات الدراسة متساوية أم أن هناك وسيطاً واحداً على الأقل من تلك الأوساط يختلف عن البقية اختلافاً معنوياً.

متى يستخدم:

يستخدم اختباركروسكال - واليس اللامعلمي كبديل لاختبار التباين الأحادي (المعلمي) لمتوسطات ثلاث عينات مستقلة أو أكثر وذلك عند مخالفة اختبار F لأحد شروط استخدامه مثل كون توزيع المتغير التابع في مجتمعات الدراسة غير طبيعي، أو عدم تجانس تباينات المجتمعات أو لكون بيانات المتغير التابع رتبية.

ملاحظة: يمكن استخدام اختبار كروسكال - واليس لمقارنة وسيطي عينتين مستقلتين. ويعتبر اختبار مان - ويتني حالة خاصة من اختبار كروسكال-واليس.

شروط استخدامه:

١- العينات مستقلة بعضها عن بعض.

٢- أن تكون العينات مختارة عشوائياً من مجتمعات الدراسة.

٣- أن يكون المتغير التابع محل الدراسة رتبياً على الأقل (يعني رتبي، فتري أو نسبي).

٤- أن يكون توزيع المتغير التابع في مجتمعات الدراسة متماثلاً في الشكل.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سيتم توضيح كيفية تنفيذ اختبار كروسكال - واليس لأوساط ٣ عينات مستقلة فأكثر وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٧-٤):

البيانات في الجدول التالي تبين تكرارات درجة الرضا العام لمقدمي الخدمة بالشؤون الصحية في المدن الرئيسية بالمملكة العربية السعودية. تم أخذ عينة عشوائية من مقدمي الخدمة في كل مدينة من المدن الرئيسية وطلب منهم تحديد مستوى رضاهم العام عن خدمات الرعاية الصحية الأولية المقدمة في المملكة (فهمي، ٢٠٠٥، ٤٩٢).

جدول رقم (٧-٨)

درجة الرضا العام لمقدمي الخدمة بالشؤون الصحية في المدن الرئيسية بالمملكة العربية السعودية

درجة الرضا / المدينة	غير راضٍ تماماً (١)	غير راضٍ (٢)	غير محدد (٣)	راضٍ (٤)	راضٍ تماماً (٥)
الرياض	١	٨	١	٩	٠
جدة	٣	٩	٠	٥	٠
أبها	٠	٢	١	٦	٣
الدمام	٠	٢	٣	١٣	٠
تبوك	٠	٣	١	٨	٠

المطلوب: هل تدل البيانات على وجود اختلاف معنوي في مستويات الرضا العام لمقدمي الخدمة بمديريات الشئون الصحية عن خدمات الرعاية الصحية الأولية المقدمة في المملكة باختلاف المدن الرئيسية؟

خطوات الحل:

١- صياغة الفرضيات:

H_0: لا يوجد اختلاف معنوي بين أوساط مستوى الرضا العام لمقدمي الخدمة عن خدمات الرعاية الصحية الأولية المقدمة في المملكة باختلاف المدن الرئيسية.

H_1: يوجد اختلاف معنوي في وسيط مستوى الرضا العام لمقدمي الخدمة عن خدمات الرعاية الصحية الأولية المقدمة في المملكة في مدينتين على الأقل من المدن الرئيسية.

٢- تحديد مستوى المعنوية: وليكن $\alpha =$ ٠,٠٥.

٣- تحديـد الاختبار الإحصائي المناسـب: الاختبار الإحصائي فـي هذه الحالة هو اختبار كروسكال - واليس (Kruskal-Wallis Test).

ومن ثم تأخذ إحصائية الاختبار الشكل التالي:

عدد المجموعات

مجموع الرتب في المجموعة i

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \left[\frac{\left(\sum R_i\right)^2}{n_i} \right] - 3(N+1) \qquad N = \sum_{i=1}^{k} n_i$$

ملاحظة: ينبغي التنويه إلى:

- أن هناك صياغة مختلفة لإحصائية اختبار كروســكال - واليس H يمكنك الاطلاع عليها من خلال دليل SPSS في قائمة Help.
- أنه يتم تحوير إحصائية الاختبار H في حال وجود رتب متساوية وفي حال استخدام توزيعات احتمالية أخرى كتوزيع F وتوزيع χ^2 كتقريب لتوزيع H.

٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. ومن ثم تأخذ البيانات بعد إدخالها الشكل التالي:

شكل رقم (٧-٧)
شكل البيانات بعد إدخالها بواسطة SPSS

تطبيق١٢.sav [DataSet1] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

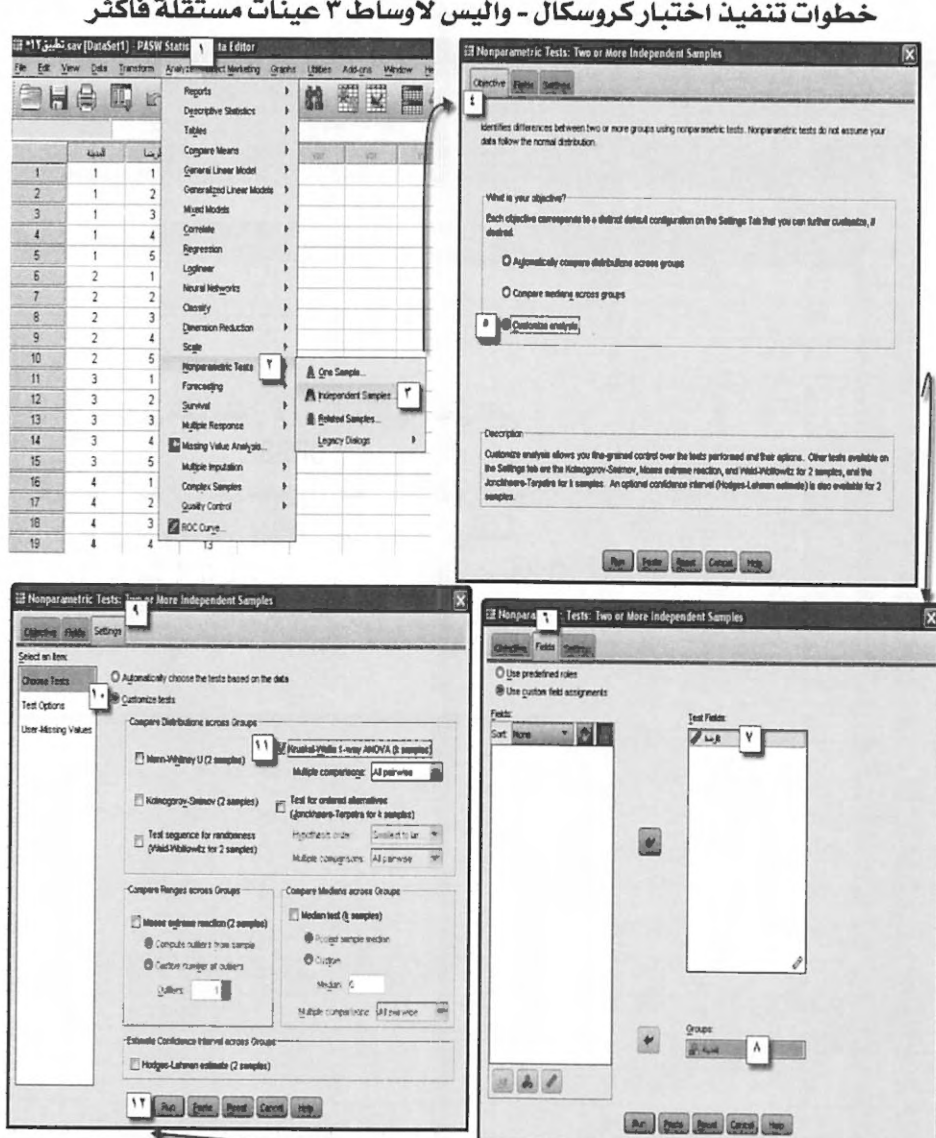
	المدينة	الرضا	التكرار	var	var	var	var	var	var
1	1	1	1						
2	1	2	8						
3	1	3	1						
4	1	4	9						
5	1	5	0						
6	2	1	3						
7	2	2	9						
8	2	3	0						
9	2	4	5						
10	2	5	0						
11	3	1	0						
12	3	2	2						
13	3	3	1						
14	3	4	6						
15	3	5	3						
16	4	1	0						
17	4	2	2						
18	4	3	3						
19	4	4	13						
20	4	5	0						
21	5	1	0						
22	5	2	3						
23	5	3	1						
24	5	4	8						
25	5	5	0						

قبل البدء في تنفيذ اختبار كروسكال-واليس، ينبغي اخبار SPSS
أن متغير "التكرار" بأنه متغير وزن يعكس عدد الذين أجابوا "موافق تماما " مثلا، وذلك من خلال الذهاب إلى قائمة
Data--->Weight Cases--->Weight Cases by--->
التكرار---> OK

بعد ذلك تتبع الخطوات التسلسلية من ١ إلى ١٢ كما هو موضح بالشكل (٧-٨) لتنفيذ الاختبار.

شكل رقم (٧-٨)

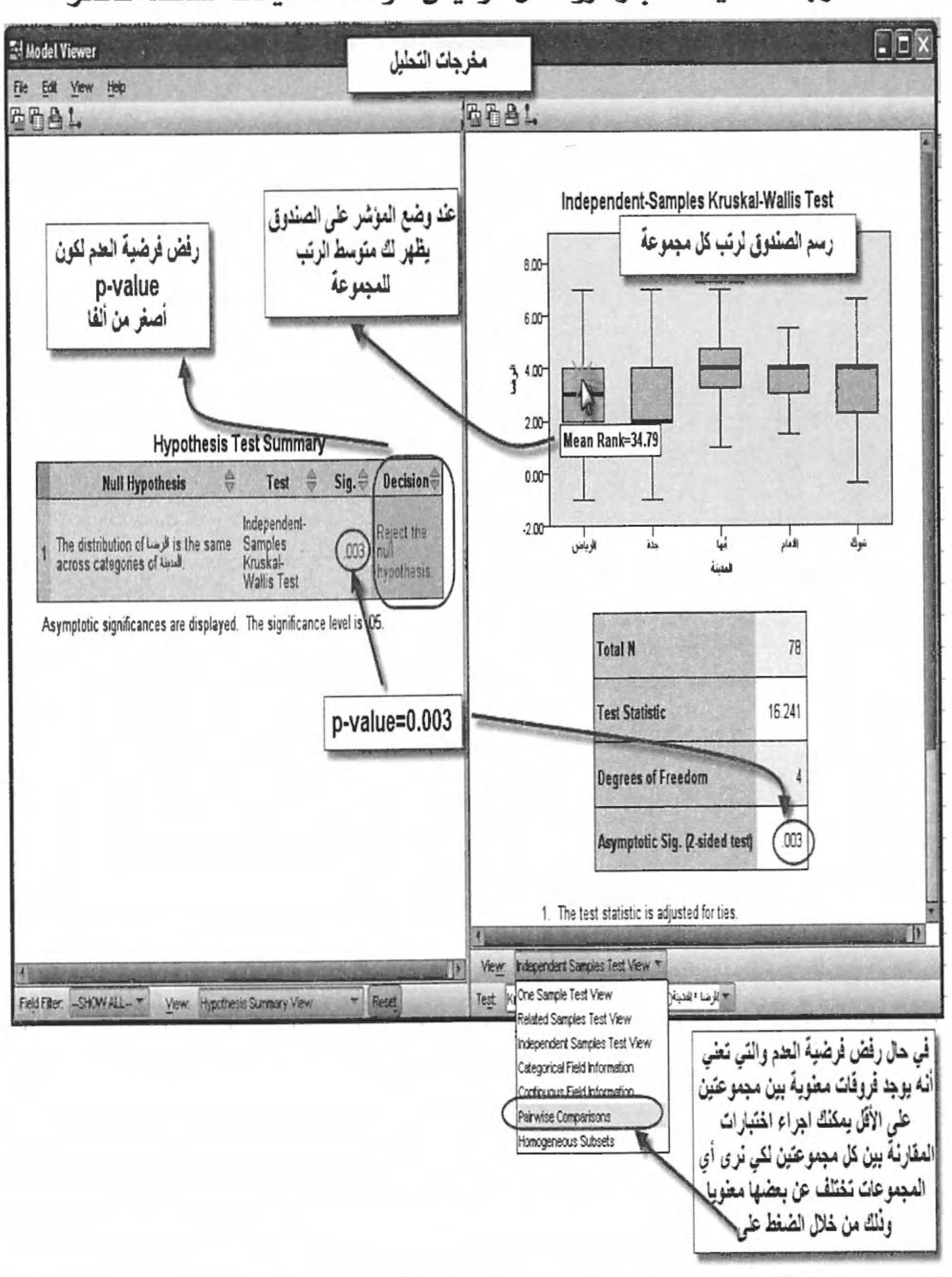
خطوات تنفيذ اختبار كروسكال - واليس لأوساط ٣ عينات مستقلة فأكثر



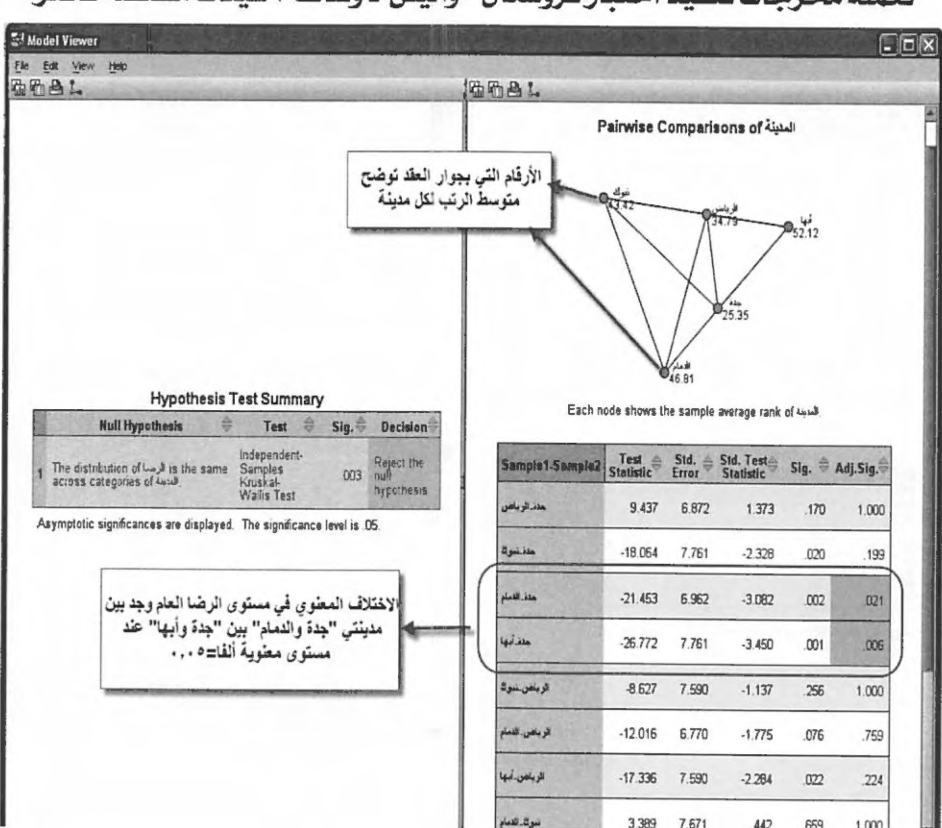
وبعـــد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output، نقوم بالنقر عليه مرتين متتاليتين للحصول على التالي:

شكل رقم (٧-٩أ)

مخرجات تنفيذ اختبار كروسكال - واليس لأوساط ٣ عينات مستقلة فأكثر



شكل رقم (٧-٩ ب)
تكملة مخرجات تنفيذ اختبار كروسكال - واليس لأوساط ٣ عينات مستقلة فأكثر



	Null Hypothesis	Test	Sig.	Decision
1	The distribution of الرضا is the same across categories of المدينة.	Independent-Samples Kruskal-Wallis Test	.003	Reject the null hypothesis

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
جدة-الرياض	9.437	6.872	1.373	.170	1.000
جدة-تبوك	-18.064	7.761	-2.328	.020	.199
جدة-الدمام	-21.453	6.962	-3.082	.002	.021
جدة-أبها	-26.772	7.761	-3.450	.001	.006
الرياض-تبوك	-8.627	7.590	-1.137	.256	1.000
الرياض-الدمام	-12.016	6.770	-1.775	.076	.759
الرياض-أبها	-17.336	7.590	-2.284	.022	.224
تبوك-الدمام	3.389	7.671	.442	.659	1.000

٢- بما أن (٠,٠٥) α > (٠,٠٠٣) P-value فإن الباحث سيقوم برفض فرضية العدم. وهذا يقود إلى اســتنتاج أنه عند مستوى معنوية ٠,٠٥ = α فإن بيانات العينة تدل على وجود اختلاف معنوي بين وســيطي مســتوى الرضا العام لمقدمي الخدمة في مدينتــين على الأقل من المــدن الرئيســية، $(H(4) = 16.24,\ \text{P-value} = 0.003)$. الجدول (٧-٩) يوضح متوســط الرتب لــكل مجموعة وأحجام العينات من مقدمي الخدمة في كل مدينة من المدن الرئيســية. كما أن الشــكل (٧-٩ ب) يوضح نتائج المقارنات الثنائية البعدية لمقارنة مستوى الرضا العام لمقدمي الخدمة حسب المدينة وقــد اتضح أن هناك اختلافاً معنوياً في مســتوى الرضا العــام بين مدينتي جدة والدمام، وكذلك بين مدينتي جدة وأبها وذلك عند مستوى معنوية ٠,٠٥ = α.

جدول رقم (٧-٩)
متوسط الرتب وأحجام العينات من مقدمي الخدمة في كل مدينة من المدن الرئيسية

المدينة	الرياض	جدة	أبها	الدمام	تبوك
متوسط الرتب	٣٤,٧٩	٢٥,٣٥	٥٢,١٢	٤٦,٨١	٤٣,٤٢
حجم العينة	١٩	١٧	١٢	١٨	١٢

تمرين تطبيقي (٧-٦):

في سـؤال بحثي لمعرفة ما إذا كان لمؤهل العلمي تأثير على الآراء حول عمل المرأة، قام باحث باسـتطلاع آراء مجموعة مكونة من (١٣) شخصاً بمؤهلات علمية مختلفة (متوسط، ثانوي، جامعي) تم اختيارهم بشكل عشوائي. الجدول التالي يحتوي المؤهل ودرجات الرأي حول عمل المرأة (الدرجة الأقل تعني عدم تأييد عمل المرأة):

جدول رقم (٧-١٠)
استطلاع الآراء حول عمل المرأة حسب المؤهل العلمي

المؤهل	درجات الرأي حول عمل المرأة				
متوسط	٤٠	٣٥	٣٢	٢٣	٢٠
ثانوي	٤٢	٢٥	٢٤	١٨	--
جامعي	٤٨	٤٦	٤١	٣٣	--

المطلوب: عند مسـتوى معنوية $\alpha =$ ٠,٠٥، هل تـدل البيانات على وجود اختلاف معنوي في الرأي حول عمل المرأة باختلاف المؤهل العلمي؟

تمرين تطبيقي (٧-٧):

الجدول التالي يحتوي درجات عينة عشـوائية من طلاب الصف الثالث الثانوي في اختبار الـ SAT موزعة حسب فئات معدل أدائهم الدراسي GPA.

جدول رقم (٧-١١)

توزيع درجات عينة عشوائية من طلاب الصف الثالث الثانوي في اختبار الـ SAT حسب معدل أداءهم الدراسي GPA

٠ – ٢,٩٩	٣,٠ – ٣,٥٠	٣,٥١ – ٤,٠	٠ – ٢,٩٩	٣,٠ – ٣,٥٠	٣,٥١ – ٤,٠
١٠٤٠	٨٧٠	١٢١٠	١٢٢٠	١٢٧٠	١٢٢٠
٩٦٠	١٠١٠	١٠٣٠	٨٥٠	١٢٠٠	١١٦٠
٩٢٠	٧٦٠	١٠٢٠	٨٨٠	٩٧٠	٩٥٠
١٢٦٠	٨٣٠	١٤٠٠	٩٢٠	٧٢٠	١٠٧٠
١٠٦٠	٩١٠	١١٧٠	٩١٠	١٠٧٠	١١١٠
١٠٩٠	١٢٣٠	١٣٥٠	١٢٦٠	٨٦٠	١١٥٠
١٠٥٠	١١١٠	١٠٧٠	٨٣٠	٩٨٠	١٤٤٠
٨٨٠	٩٧٠	١٣٢٠	٨١٠	٨٠٠	١٢٨٠
٧٥٠	١٠٩٠	١٢٠٠	٨٢٠	٩٠٠	١٠٢٠
٩٦٠	١٣٠٠	١٠٧٠	١١٧٠	١٣٣٠	١٠٨٠
١٠٩٠	–	١٢٤٠	٨١٠	١١٤٠	١١٢٠

المطلوب: عند مستوى معنوية ٠,٠٥ = α، هل تدل البيانات على وجود اختلاف معنوي في درجة اختبار الـ SAT باختلاف معدل الأداء الدراسي GPA؟

الفصل الثامن
مقاييس الارتباط بين المتغيرات Measures of Association

مقدمة:

إن مقاييـــس الارتباط بين المتغيـــرات تعتبر من أكثر أســـاليب التحليل الإحصائي المستخدمة في الدراسات والبحوث العلمية وفي مختلف المجالات العملية. ففي الحياة اليوميـــة نرى كثيراً من الظواهر والمتغيرات التي تتغيـــر قيمها وفقاً للتغير في ظواهر ومتغيرات أخرى سواء في نفس الاتجاه أو في اتجاه معاكس ووفق نمط معين وبدرجات متفاوتة. وهذا يقودنا إلى التســـاؤل حول نوعية العلاقة والارتباط بين تلك المتغيرات. ومن أمثلة ذلك، العلاقة بين العرض والطلب على السلع، دخل الأسر وإنفاقها، المستوى الاقتصادي ومعدل الجرائم، معدل ساعات التدريب للموظف وإنتاجيته ...إلخ. وتهدف دراســـة الارتباط بين المتغيرات إلى وصف نوعية واتجاه وقوة العلاقة بينها بحيث يتم الاســـتفادة منها في اتخاذ القرارات ورسم الخطط والسياسات وبناء علاقات رياضية تربـــط بين تلك المتغيرات يســـتفاد منها في عمليات الوصـــف والتنبؤ بقيم بعض تلك المتغيـــرات بالاعتماد على قيم متغيرات أخرى. إن مقاييس الارتباط بين المتغيرات لها تطبيقات عديدة وتفيد في التالي:

١- تحديد درجة أو قوة الارتباط بين المتغيرات وبيان ما إذا كانت قوية أو متوسطة أو ضعيفة أو منعدمة.

٢- تحديد اتجاه العلاقة بين المتغيرات بمعنى بيان ما إذا كانت طردية أو عكسية.

٣- دراسة وتحليل العلاقات السببية بين المتغيرات.

٤- قياس الصدق والثبات والتحليل العاملي وتحليل الانحدار وكثير من أساليب التحليل الإحصائي الأخرى.

ولقد كان فرانســـس قالتون Francis Galton هو أول من بدأ مفهوم الارتباط بين المتغيـــرات في أواخر القرن الثامن عشـــر الميلادي. ومن ذلـــك الحين وحتى الآن تم تطويـــر وتصميم العديد مـــن مقاييس الارتباط آلتي تعالج مختلـــف الظروف المتعلقة بطبيعة البيانات ومستويات قياس المتغيرات وعدد المتغيرات وغيرها من الظروف التي تحـــدد نوع المقياس والحاجة إليه. إن المتصفح للأدبيـــات المتعلقة بمقاييس الارتباط

بين المتغيرات سـيجد أنواعاً كثيرة ومتعددة. والجدول (٨-١) يقدم عرضاً لمعظم تلك المقاييس ويعتبر دليلاً إرشادياً يمكن بناء عليه اختيار المقياس المناسب بناء على نوع قياس المتغيرات أو الهدف من دراسة العلاقة (Glass & Hopkins (1996), Conover(1999 وSPSS® V.20. لذا سيتم الاقتصار في هذا الكتاب على اختيار بعض تلك المقاييس والتي تعتبر من أكثر مقاييس الارتباط اسـتخداماً وشـيوعاً في الدراسات والبحوث الميدانيـة في كثير من مجالات العلوم كالإدارية والإنسـانية والاجتماعية والاقتصادية وغيرها.

جدول رقم (٨-١)

دليل إرشادي لاختيار المقياس الأنسب من بين مقاييس الارتباط بين المتغيرات

اسم معامل الارتباط	الصيغة الرياضية	مستوى قياس المتغيرات	قيمته العددية	إيضاحات
بيرسون Pearson	$r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$	كلا المتغيرين من المستوى الفتري أو النسبي	$-1 \le r \le +1$	- يوجد ارتباط خطي قوي بين المتغيرين كلما اقتربت قيمة r من الواحد الصحيح سواء في الاتجاه الموجب أوالسالب وتنعدم العلاقة الخطية كلما اقتربت قيمة r من الصفر. - تتأثر قيمة معامل بيرسون للارتباط بشكل كبير بالقيم الشاذة فقد تكون قيمة r قريبة من الواحد علماً أنه لا يوجد ارتباط خطي بين المتغيرين أو العكس - انظر الشكل (٩-٣) في الفصل التاسع -، لذا ينبغي استخدام الشكل الانتشاري لتمثيل العلاقة بين المتغيرين بيانياً قبل حساب الارتباط للتعرف على طبيعة توزيع البيانات لنرى ما إذا كان هناك قيم شاذة وطريقة توزيع البيانات في مجملها.
سبيرمان Spearman	$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$	كلا المتغيرين من المستوى الرتبي	$-1 \le r_s \le +1$	- يفسر بنفس طريقة معامل بيرسون. - يمكن استخدامه أيضاً مع المتغيرات من المستوى الفتري أو النسبي.
جاما Gamma	$\gamma = \frac{C-D}{C+D}$	كلا المتغيرين من المستوى الرتبي	$-1 \le \gamma \le +1$	- يفسر بنفس طريقة سبيرمان. ويفضل على سبيرمان وكندال تاو kendall's Tau إذا كان هناك العديد من القيم أو المشاهدات المتساوية Tied data. - هناك معاملات ارتباط أخرى تستخدم في حالة البيانات الرتبية وهي تقريباً مشابهة إلى حد ما معاملي سبيرمان وجاما وتتراوح قيمها بين (-١) و(+١). ومن هذه المقاييس معامل سومرز دي Somers' d ومعاملات كندال تاو Kendkasll's Tau-a, Kendkasll's Tau-b, Kendkasll's Tau-c. ولكن يعتبر معاملا سبيرمان وجاما من أكثر مقاييس الارتباط بين المتغيرات الرتبية استخداماً وانتشاراً.

تابع جدول رقم (٨-١).

اسم معامل الارتباط	الصيغة الرياضية	مستوى قياس المتغيرات	قيمته العددية	إيضاحات
مربع كاي χ^2	$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$	كلا المتغيرين من المستوى الاسمي	$0 \le \chi^2 \le +\infty$	- قيمته دائماً موجبة ويستخدم في حالة جداول الاقتران* لاختبار استقلالية (عدم ارتباط) متغيرين إذا كان عدد الصفوف والأعمدة ٢ فأكثر. ولا يشترط أن يكون الجدول مربعاً، أي عدد الصفوف يساوي عدد الأعمدة. وإذا كان عدد الصفوف يساوي ١ وعدد الأعمدة ٢ فأكثر فإنه في هذه الحالة يسمى مربع كاي لجودة التوفيق أو حسن المطابقة.
معامل التوافق Contingency coefficient	$cc = \sqrt{\frac{\chi^2}{\chi^2 + n}}$	كلا المتغيرين من المستوى الاسمي	$0 \le cc \le \sqrt{\frac{k}{k-1}}$ حيث $k = \min(r,c)$	- يستخدم في حالة جداول الاقتران* لأي عدد من الصفوف والأعمدة، وهو يعتمد كما هو واضح من الصيغة الرياضية له على مربع كاي. والميزة التي يتميز بها معامل التوافق عن مربع كاي تتمثل في أنه أسهل في عملية تفسير العلاقة بين المتغيرين محل الدراسة. ويصعب الوصول إلى القيمة واحد، حيث إن الاقتراب من القيمة واحد تتطلب عدداً كبيراً من الصفوف والأعمدة.
معامل فاي Phi	$\phi = \sqrt{\frac{\chi^2}{n}}$	كلا المتغيرين من المستوى الاسمي	$0 \le \phi \le +1$	- يستخدم في حالة جداول الاقتران* من أي حجم وينبغي ملاحظة أنه في حالة الجداول من الحجم 2x2 فإن قيمة وإشارة معامل فاي تساوي معامل ارتباط بيرسون. كما ينبغي ملاحظة أن قيمة ϕ تتراوح بين الصفر والواحد الصحيح فقط في حالة الجداول من الرتبة 2x2، 2xc، وrx2 ولكن تفقد ϕ هذه الميزة عندما يكون عدد كلٍّ من الصفوف والأعمدة أكبر من ٢.
معامل كرامير Cramer's V	$V = \sqrt{\frac{\chi^2}{n \times (\min(r,c) - 1)}}$	كلا المتغيرين من المستوى الاسمي	$0 \le V \le +1$	- يستخدم في حالة جداول الاقتران* لأي عدد من الصفوف والأعمدة. ونلاحظ أنه عبارة عن تحوير بسيط لمعامل ϕ للتغلب على الصفة السلبية لمعامل ϕ عندما يكون كلٌّ من عدد الصفوف والأعمدة أكبر من ٢. ومن ثم فإن قيمة V تتراوح دائماً بين الصفر والواحد الصحيح لأي عدد من الصفوف والأعمدة. ولهذا فإن معامل كرامير يفضل على معاملي التوافق وفاي.

تابع جدول رقم (٨-١).

اسم معامل الارتباط	الصيغة الرياضية	مستوى قياس المتغيرات	قيمته العددية	إيضاحات
معامل لامبدا Lambda	$\lambda=\frac{\sum_{i=1}^{r}\max(o_{ij})+\sum_{j=1}^{c}\max(o_{ij})-\max(r_i)-\max(c_j)}{2n-\max(r_i)-\max(c_j)}$	كلا المتغيرين من المستوى الاسمي	$0 \le \lambda \le +1$	- لامبدا هو معامل ارتباط يمثل نسبة تقليص مقدار الخطأ في التنبؤ بالمتغير التابع باستخدام قيم المتغير المستقل. فالقيمة ١ تعني أن المتغير المستقل يتنبأ بالمتغير التابع بالشكل التام والعكس أي أن القيمة٠ تعني أنه لا يصلح أبداً التنبؤ بقيمة المتغير التابع من المستقل لأنه لا يوجد ارتباط بين المتغيرين. وتجدر الإشارة إلى أن الصيغة الرياضية السابقة لمعامل لامبدا هي صيغة معامل لامبدا المتماثل. أي أن كون المتغير الأول مستقلاً والثاني تابعاً أو العكس، سيعطي نفس القيمة العددية مع اختلاف تفسير النتائج. كما تجدر الإشارة إلى أن برنامج SPSS يعطي ثلاث قيم لمعامل لامبدا. القيمة الأولى لا تميز بين أي المتغيرين مستقل وأيهما تابع (مقياس متماثل symmetric)، أما القيمتان الثانية والثالثة فهما تميزان بين: أي المتغيرين تابع وأيهما مستقل (مقياس غير متماثل asymetric). - معامل عدم التأكد Uncertainty coefficient يماثل معامل لامبدا في الهدف من استخدامه والتفسير، لكنه يختلف عنه في الصيغة الرياضية.
إيتا Eta	$\eta=\sqrt{1-\frac{SS_{within}}{SS_{total}}}$	أحد المتغيرين اسمي والآخر فتري أو نسبي	$0 \le \eta \le +1$	- يستخدم لحساب قوة الارتباط بين متغيرين في حالة العلاقات غير الخطية nonlinear بين المتغيرات. ومربع إيتا يشير إلى أي درجة يستطيع أحد المتغيرين (المستقل - ذي المستوى الاسمي) التنبؤ بقيم المتغير الآخر (التابع - ذي المستوى الفتري أو النسبي). - هناك معاملات ارتباط أخرى تستخدم في هذه الحالة مثل معامل ارتباط السلسلتان الثنائي والذي يعتبر حالة خاصة من معامل ارتباط بيرسون. - Point biserial conrrelation ومعامل إيبسلون (ε).

تابع جدول رقم (٨-١).

اسم معامل الارتباط	الصيغة الرياضية	مستوى قياس المتغيرات	قيمته العددية	إيضاحات
مقاييس أخرى	- مقياس كابا للاتفاق Kappa measure of agreement لقياس الاتفاق بين تقييمات مقيمين مختلفين لنفس الموضوع. - مقياس تقدير المخاطرة Risk assessment يستخدم في حالة الجداول من الحجم 2x2 لقياس العلاقة ما بين عامل (متغير) معين وحدوث حادثة معينة. - مقياس أو اختبار ماكنيمر McNemar يستخدم للجداول 2X2 لاختبار ما إذا كان هناك حدوث تغير معنوي في الاستجابة بعد تنفيذ تجربة معينة وذلك من خلال مقارنة الاستجابات قبل وبعد التجربة. - مقياس أو اختبار كوكران ومانتل - هانزل Cochran's and Mantel-Haenzel statistics وهو تعميم لـ «ماكنيمر» أي يختبر العلاقة ما بين متغيرين ثنائيين Dichotomous عند كل المستويات المختلفة من المتغير الثالث (متغير التحكم).			

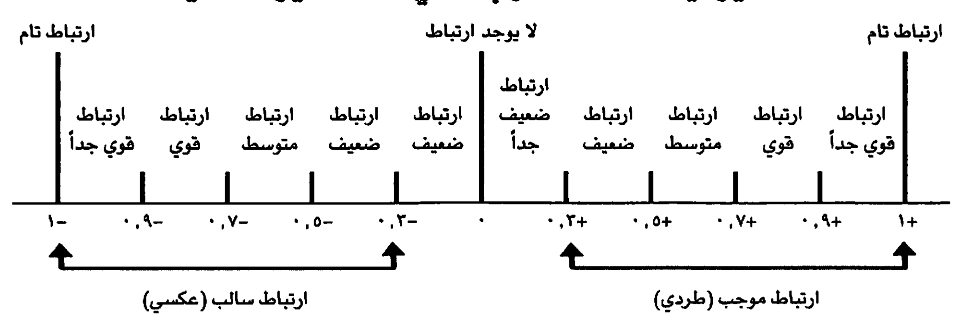
* جدول الاقتران Contingency table (Cross table): عبارة عن جدول على هيئة مصفوفة مكونة من صفوف وأعمدة لعرض التوزيع التكراري للمتغيرات.

- r وc ترمزان لعدد الصوف والأعمدة في جدول الاقتران على التوالي.
- o وe ترمزان لعدد التكرارات المشاهدة والتكرارت المتوقعة لحدث ما على التوالي.
- C وهو الحرف الأول من Concordant pair وتعني الأزواج المرتبة المتسقة. فيقال لمشاهدة من البيانات أنها متسقة في قيمها إذا كانت قيم المتغيرين لتلك المشاهدة أعلى من قيم المتغيرين المناظرة لمشاهدة أخرى. فمثلاً عند مقارنة قيم المشاهدة (٤،٥) مع المشاهدة (٢،٣) فإننا نلاحظ أن القيمة ٥ من المشاهدة الأولى أكبر من ٣ في المشاهدة الثانية وكذلك فإن القيمة ٤ في المشاهدة الأولى أكبر من ٢ في المشاهدة الثانية.
- D وهو أول حرف من Discordant pair وتعني الأزواج المرتبة المتناقضة. فيقال لمشاهدة من البيانات أنها متناقضة في قيمها إذا كانت إحدى قيم المتغيرين لتلك المشاهدة أعلى من نظيرتها لمشاهدة أخرى بينما القيمة الأخرى في المشاهدة الأولى أقل من نظيرتها في المشاهدة الثانية. فمثلاً عند مقارنة قيم المشاهدة (٤،٥) مع المشاهدة (٢،٧) فإننا نلاحظ أن القيمة ٥ من المشاهدة الأولى أصغر من ٧ في المشاهدة الثانية في حين أن القيمة ٤ في المشاهدة الأولى أكبر من ٢ في المشاهدة الثانية.
- n الحجم الكلي للعينة.
- SSwithin ترمز للمجموع الكلي لمجموع مربعات انحرافات القيم عن وسطها الحسابي داخل الجموعات.
- SStotal وترمز للمجموع الكلي لمربع انحرافات القيم عن الوسط الحسابي الكلي للبيانات.

تفسير قيمة معاملات الارتباط الكمية والنوعية:

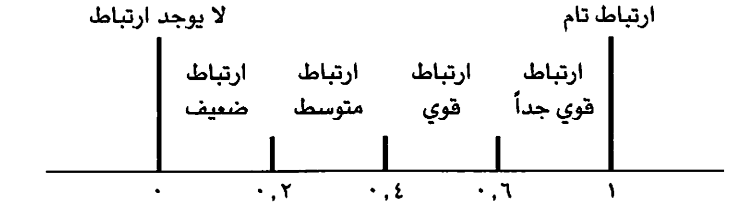
لقد وضع الباحثون والمختصون حدوداً إرشادية عامة rule of thumb يمكن للباحث الاستعانة بها في تفسير قيمة معامل الارتباط الخطي، وهي قد تختلف حسب طبيعة الدراسة ومجالها (زايد، ٢٠٠٧). فمثلاً تعتبر قيمة معامل الارتباط التي تساوي ٠,٦٠ ارتباطاً متوسطاً في أحد مجالات العلوم كالعلوم الاجتماعية، لكنها تعتبر ضعيفة في مجالات أخرى كالمجال الاقتصادي. ويمكن الاسترشاد بالتجزيء المقترح في الشكل (٨-١) لتفسير قيم معامل الارتباط في حالة المتغيرات الكمية.

شكل رقم (٨-١)
تفسير قيمة معاملات الارتباط في حالة المتغيرات الكمية



أما في حالة معاملات الارتباط للمتغيرات الاسمية كمعامل التوافق، فاي، كرامير وغيرها فإنه يمكن تفسير قيمها التي تقع بين الصفر والواحد بطريقة تختلف عن معاملات الارتباط للمتغيرات الكمية. ويمكن الاسترشاد بالتجزيء المقترح في الشكل (٨-٢) لتفسير معامل الارتباط في حالة المتغيرات الاسمية (زايد، ٢٠٠٧).

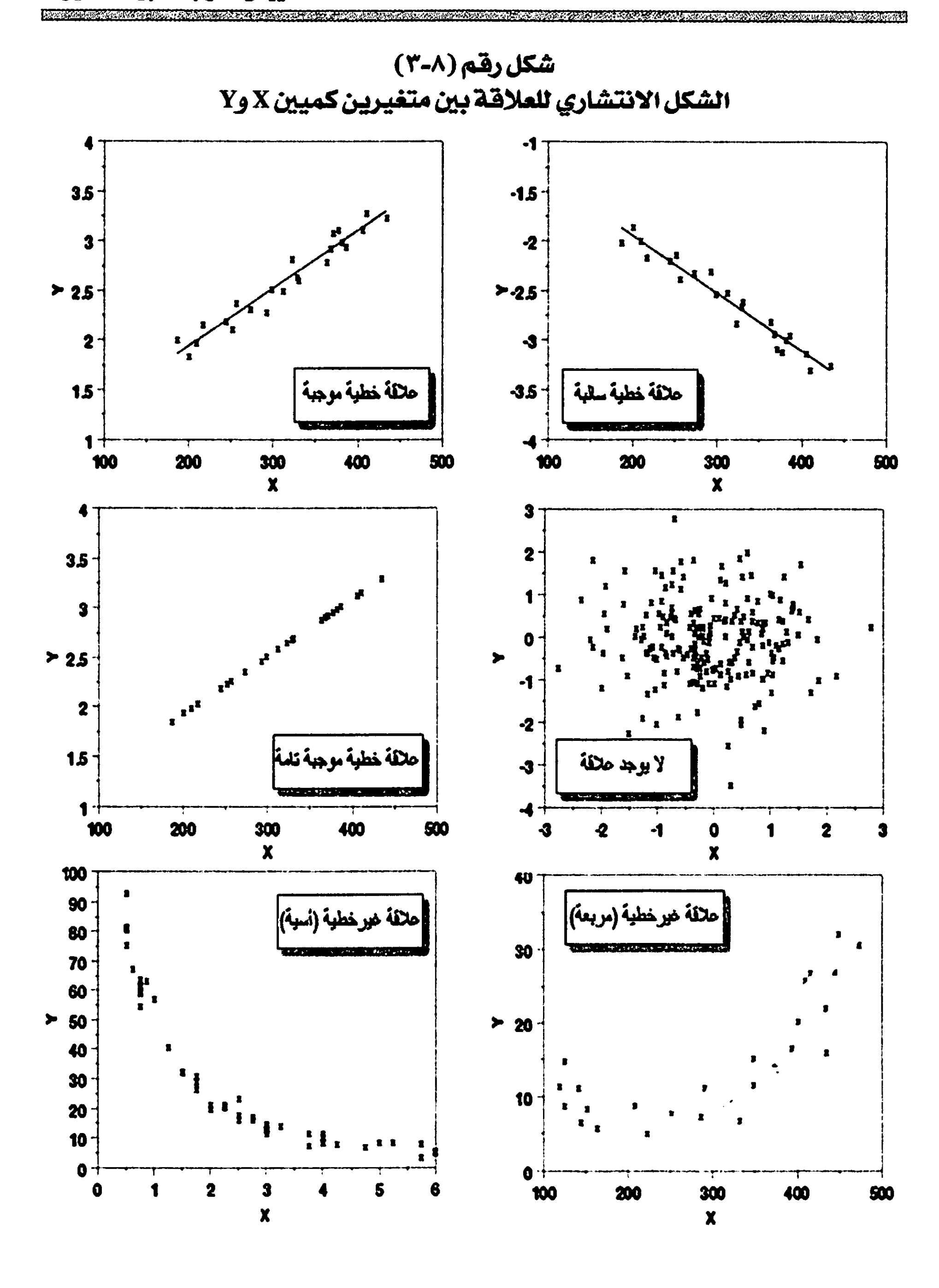
شكل رقم (٨-٢)
تفسير قيمة معاملات الارتباط في حالة المتغيرات الاسمية



العلاقة بين متغيرين كميين:

إذا كان لدينـــا متغيـــران كميان X وY فإن العلاقة بينهمـــا قد تكون خطية، أو غير خطية، أو منعدمة. كما أن العلاقة بين متغيرين كميين قد تكون موجبة (طردية) أي أن المتغيرين تتزايد أو تتناقص قيمهما في نفس الاتجاه أو قد تكون علاقة سالبة (عكسية) بمعنى أن قيم المتغيرين تتحركان في اتجاهين متعاكســين، بمعنى أنه إذا تزايدت قيم أحــد المتغيرين مثلاً، فإن قيــم المتغير الثاني تتناقص والعكــس صحيح. وربما تكون العلاقة بين المتغيرين طردية في فترة معينة وعكسية في فترة أخرى والعكس صحيح. وقبل البدء في دراســة علاقة الارتباط بــين المتغيرات يوصى بتمثيلهما بيانياً للتعرف على شكل توزيع البيانات بصرياً حيث يمكن للباحث أو محلل البيانات أن يتعرف على طبيعــة العلاقة بين المتغيرات بيانياً كأن تكون خطية، غيــر خطية، أومنعدمة وكذلك اتجاه وقوة العلاقة، كما يمكنه رؤية ما إذا كان هناك قيم شــاذة أو متطرفة عن بقية البيانات. والشــكل (٨-٣) يوضح بيانياً من خلال الأشــكال الانتشارية scatter plots بعضاً من أنماط العلاقات الارتباطية بين متغيرين كميين.

شكل رقم (٨-٣)
الشكل الانتشاري للعلاقة بين متغيرين كميين X وY



الارتباط والسببية Correlation and Causality:

من الأخطاء الشــائعة بين كثير من الباحثين أنه بمجــرد وجود علاقة ارتباط بين متغيرين يعني أن التغير في أحدهما يسبب التغير في الآخر. إن وجود ارتباط خطي (أو غير خطي) بين متغيرين لا يعني بالضرورة أن التغير في قيم أحدهما هو الســبب فــي التغير في قيم المتغير الآخر، وإنما هذا يعني أن هناك توافقاً أو تزامناً في ظهور قيم المتغيرين بطريقة معينة، فمثلاً القيم الكبيرة من المتغير الأول تتزامن في الحدوث مع القيم الكبيرة من المتغير الثاني والصغيرة مع الصغيرة (علاقة طردية) أو العكس – أي ظهــور القيم الكبيرة مــن المتغير الأول مع القيم الصغيرة من المتغير الثاني (علاقة عكسية).

إن معامــلات الارتباط تبين لنا فقط مــا إذا كان هناك علاقة خطية بين متغيرين ومدى قــوة واتجاه الارتباط، ولكنها لا تســتطيع إخبارنا عمــا إذا كانت تلك العلاقة ســببية أو لا. وقد ذكر Madsen (2011) أن المعرفة بطبيعة مجال الدراســة يساعد فــي الحكم على ســببية العلاقة بين المتغيرات، وأشــار إلى أن الكثيــر من المتغيرات فــي العلــوم الاجتماعية والطبيعية تتزايــد مع الزمن. ولهذا فإن كثير من دراســات الارتباط بين المتغيرات وتحديداً الدراســات غير التجريبيــة تظهر وجود ارتباط بين المتغيرات، وهذا بســبب أن تلك المتغيرات تتغير في نفس الاتجاه أو في اتجاه معاكس مع الزمن (Agresti and Finlay, 2009). فمثلاً في دراسة لعلاقة الارتباط بين طول الطــلاب وتحصيلهم في مادة الرياضيات أخذت عينة عشــوائية وتبين وجود ارتباط خطــي طردي بين الطول والتحصيل للطلاب، فهل هذا يعني أن الطالب الأطول يكون تحصيله أفضل من الطالب الأقصر؟! ليس بالضرورة صحيحاً وليس من المنطق قبول مثل هذا الاســتنتاج. إذن ما الســبب في ظهور ذلك الارتبــاط الخطي الطردي بين المتغيرين؟ إن السبب يعزى إلى وجود متغير آخر وهو متغير العمر والذي يعتبر متغيراً كامناً Lurking variable، وهو الســبب في تزايد قيم المتغيرين معاً، فمع نمو الطالب يزداد طوله وتحصيله وهذا يقود إلى أن العلاقة بين متغيري الطول والتحصيل علاقة زائفة Spurious relationship ليست سببية.

متــى يمكن القول بأن علاقة الارتباط بين متغيري X وY هي علاقة ســببية؟ إذا كان هناك علاقة بين المتغيرين X وY فإنه يمكن النظر لها على أنها سببية إذا تحققت الشروط الثلاثة التالية (Agresti & Finlay, 2009):

١- أن يكون هناك علاقة ارتباط قوية نسبياً بين المتغيرين X وY.

٢- أن يكون أحد المتغيرين وليكن X (المتغير المستقل أو السبب) يسبق المتغير الآخر Y(المتغير التابع أو النتيجة) في الحدوث. فمثلاً في العلاقة بين العمر والتحصيل العلمي، نجد أن العمر يسبق ظهور التحصيل وهذا ترتيب زمني طبيعي. وعادة في الدراسات غير التجريبية أو الدراسات الوصفية يصعب تحديد الترتيب الزمني للمتغيرات، وكمثال على ذلك العلاقة بين الحالة الاقتصادية للزوج والحالة الاقتصادية للزوجة.

٣- استبعاد أو ضبط أي متغيرات أخرى بديلة يعتقد بأنها تؤثر على تفسير العلاقة بين المتغيرين X وY. وتعتبر هذه الخاصية أصعب من سابقتيها إذ يصعب تحديدها. فكثيراً ما تظهر الدراسات في عدة مجالات مختلفة بأن هناك علاقات ارتباط بين متغيرات معينة، ولكنها في الحقيقة قد تكون زائفة (غير حقيقية) وإنما هي نتاج لعوامل أخرى. فمثلاً في الدراسات الطبية يقال بأن هناك علاقة بين شرب القهوة واحتمالية الإصابة بالنوبات القلبية. ولكن بعد أخذ بعض المتغيرات بالاعتبار وضبطها مثل نوعية الوظيفة، بلد المعيشة، مستوى التوتر لدى الأفراد وغير ذلك تبين أن تلك العلاقة ضعفت بشكل كبير وتلاشت.

وبالنظر إلى العلاقة بين التدخين والإصابة بمرض سرطان الرئة، فقد ذكر Agresti & Finlay (2009) بأنه يمكن اعتبارها الآن سببية لاستيفائها الشروط الثلاثة السابقة. فقد وجد أن هناك علاقة قوية إلى حد ما بين المتغيرين (الشرط الأول)، أيضاً هناك ترتيب زمني حيث إن الإصابة بسرطان الرئة عادة يكون بعد التدخين لفترة معينة (الشرط الثاني)، كما أنه لم يتم تحديد تفسيرات أو متغيرات جوهرية أخرى بديلة تفسر العلاقة بين المتغيرين (الشرط الثالث). وهذه العلاقة السببية قد تم تعزيزها بالنظريات الحيوية أو الطبية التي توضح الاحتمال الكبير للإصابة بسرطان الرئة في حال التدخين.

وفي الدراسات التجريبية Experimental studies يمكن استنتاج أن علاقة الارتباط إن وجدت بين المتغيرين X وY هي علاقة سببية؛ لأن الشروط الثلاثة السابقة يتم تحقيقها عادة في مثل تلك الدراسات.

والجدول (٨-٢) يوضح بعضاً من أنواع العلاقات بين المتغيرين X1 (متغير مستقل) وY (متغير تابع) واحتمالات تأثير المتغير الثالث X2 على العلاقة بينهما.

جدول رقم (٨-٢)
أنواع العلاقات بين متغيرين X1، وY في ظل وجود متغير ثالث X2

مخطط بياني	اسم العلاقة	ماذا يحدث بعد إزالة (أو ضبط) تأثير المتغير X2	مثال توضيحي
$X_2 \to X_1$، $X_2 \to Y$	علاقة زائفة بين المتغيرين X_1 و Y	تتلاشى العلاقة بين X_1 و Y	العلاقة بين التحصيل الدراسي للطالب (Y) وطوله (X_1) هي علاقة زائفة والسبب في ذلك هو وجود متغير ثالث وهو العمر (X_2) يؤثر في المتغيرين Y وX1. فنلاحظ أن المتغيرين Y وX1 هما نتيجة مباشرة للمتغير X_2 وهذا أدى إلى وجود توافق أو ارتباط بين قيم المتغيرين Y وX_1.
$X_1 \to X_2 \to Y$	علاقة السلسلة: X_2 متغير وسيط، وX_1 يؤثر بشكل غير مباشر على Y	تتلاشى العلاقة بين X_1 و Y	تأثير التعليم (X_1) على المستوى الصحي للفرد (Y) تأثير غير مباشر من خلال متغير ثالث مثل الدخل (X_2) والذي يعتبر تأثيره مباشر على المستوى الصحي للفرد.
$X_1 \to Y$، $X_2 \downarrow$	تفاعلية	العلاقة بين X_1 وY تتغير مع تغير مستويات المتغير X_2	العلاقة بين درجة استفادة المتدرب من التدريب (Y) ونوع المدرب (X_1) تتغير بتغير طريقة التدريب المستخدمة (X_2).
$X_2 \to Y$، $X_1 \to Y$	تأثيرات متعددة	العلاقة بين X1 وY لا تتغير	العلاقة بين إنتاجية الموظف (Y) ومؤهله الأكاديمي (X_1) وساعات التدريب (X_2).
$X_1 \to Y$، $X_1 \to X_2 \to Y$	التأثير المباشر وغير المباشر للمتغير X_1 على Y	العلاقة بين المتغير X_1 وY تتغير ولكنها لا تتلاشى	درجة ضغوط العمل على الموظف (X_1) تؤثر بشكل مباشر على مستوى إنتاجيته (Y) كما يؤثر بشكل غير مباشر على إنتاجيته من خلال المستوى الصحي للموظف (X_2).

المصدر: بتصرف من Agresti & Finlay (2009).

مقاييس الارتباط المعلمية Parametric Correlation Measures:

مقدمة:

إن هذه المقاييس تفترض شـروطاً معينة حول المتغيرات وتوزيعها الاحتمالي، وفيما يلي سـيتم تقديم أكثر تلك المقاييس اسـتخداماً لدراسة قوة واتجاه ومعنوية الارتباط بين متغيرين وشرحها بطريقة عملية تفاعلية باستخدام برنامج SPSS.

أولاً - معامل ارتباط بيرسون البسيط Pearson's Correlation Coefficient:

الهدف من استخدامه:

لوصـف العلاقة الخطية بـين متغيرين متصلين (فتري أو نسـبي) وتحديد قوتها واتجاهها.

ويرمز لمعامل ارتباط بيرسـون بالرمز (r). ومعامل ارتباط بيرسون مؤشر وصفي كمي تتراوح قيمته بين -١ و+١ (أي أن ١ + ≥ r ≥ ١ -) حيث الإشارة الموجبة تدل على أن المتغيرين مرتبطان طردياً والإشـارة السالبة تعني الارتباط العكسي بين المتغيرين. ويمكن الاستفادة من الشكل (٨-١) لتفسير قيمة r.

وتجدر الإشارة إلى أنه يمكن اختبار معنوية الارتباط الخطي بين متغيري الدراسة باستخدام إحصائية الاختبار التالية:

$$T = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

حيث r يرمز لمعامل ارتباط بيرسون، n حجم العينة، وT متغير عشوائي يتبع توزيع t بدرجات حرية n-2 (أي أن $T \sim t\,(n-2)$).

وإحصائية الاختبار هذه يتم استخدامها لاختبار فرضية العدم:

H_0: لا يوجد ارتباط خطي بين متغيري الدراسة (أي $H_0: \rho = 0$ حيث ρ ترمز لمعامل ارتباط بيرسون في حالة بيانات المجتمع وليس العينة).

مقابل الفرضية البديلة (فرضية الباحث):

H_1: يوجد ارتباط خطي معنوي بين المتغيرين ($H_1: \rho \neq 0$).

شروط استخدامه:

١- أن تكون العينة عشوائية وبيانات كل متغير مستقلة بعضها عن بعض.

٢- أن يكون المتغيران كميين (فتري أو نسبي).

٣- أن تكـون البيانات على هيئة أزواج مرتبـة ordered pairs حيث لكل وحدة تحليل (متـدرب أو طالب أو موظف مثلاً) توجد قيمتـان (x, y) الأولى x تنتمي للمتغير الأول والثانية y تنتمي للمتغير الثاني.

٤- أن يكون التوزيع المشترك للمتغيرين طبيعياً bivariate normal distribution.

٥- أن يكون هناك علاقة خطية بين المتغيرين، ويتم عادة التعرف على نوع العلاقة بين المتغيرين برسم الشكل الانتشـاري Scatter plot للمتغيرين قبل البدء في حساب معامل ارتباط بيرسون. وينبغي ملاحظة أن هذا الشرط سيتحقق في حال تحقق الشرط الرابع.

٦- أن يكون تباين المتغير التابع Y ثابتاً عند جميع قيم المتغير المستقل X.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سيتم توضيح كيفية تنفيذ معامل ارتباط بيرسون وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٨-١).

أراد باحـث تربـوي التعـرف على ما إذا كانـت هناك علاقة بين معـدل الطالب التحصيلـي (Y) ودرجته في اختبار القراءة والاسـتيعاب (X). حيث تم اختيار عينة عشـوائية قوامها ١٥ طالباً من طلاب السـنة الأولى الجامعية، وكانت النتائج كما هي موضحة بالجدول (٨-٣).

جدول رقم (٨-٣)

معدل الطالب التحصيلي (Y) ودرجته في اختبار القراءة والاستيعاب (X)

	الطالب														
	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥
X	٣٨	٥٤	٤٣	٤٥	٥٠	٦١	٥٧	٢٥	٣٦	٣٩	٤٨	٤٦	٤٤	٣٩	٤٨
Y	٢,١	٢,٩	٣,٠	٢,٣	٢,٤	٣,٧	٣,٢	١,٣	١,٨	٢,٥	٣,٤	٢,٦	٢,٤	٢,٥	٣,٣

والمطلوب: حســاب معامل الارتباط بين معدل الطالب التحصيلي (Y) ودرجته في اختبار القراءة والاســتيعاب (X)، واختبار معنوية هذا الارتباط عند مســتوى معنوية $\alpha = 0.01$.

خطوات الحل:

١- اســتخدم الشكل الانتشــاري لتصوير العلاقة بين المتغيرين بيانياً: ولعمل ذلك قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٩ كما هو موضح بالشكل (٨-٤) للحصول على الشكل الانتشاري.

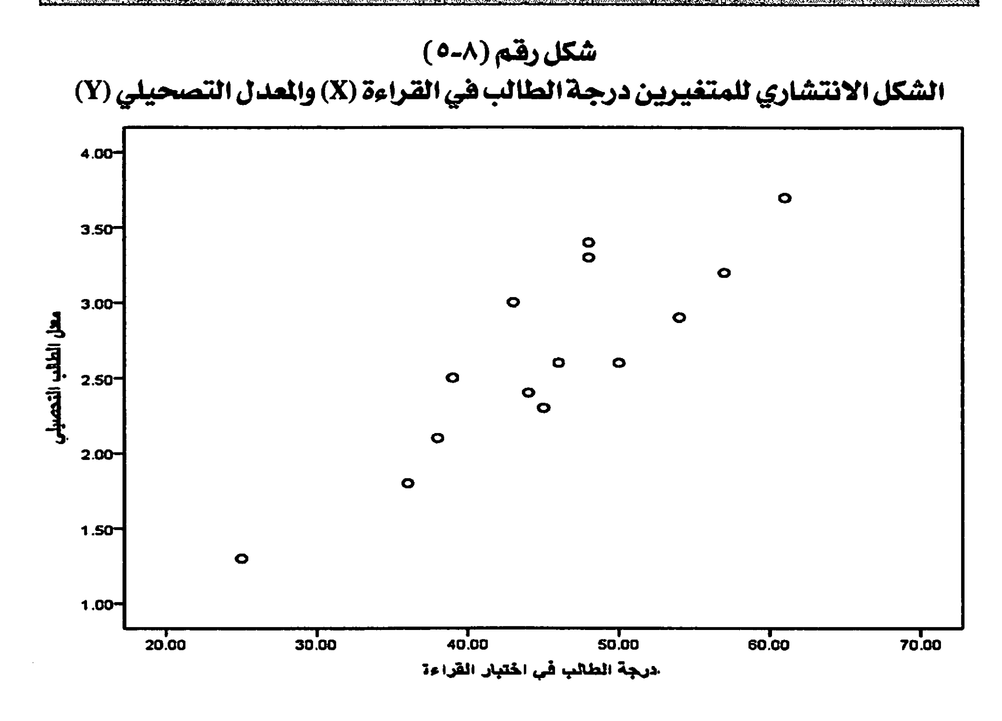
شكل رقم (٨-٤)

خطوات رسم الشكل الانتشاري للمتغيرين X وY

وبعد التنفيذ نحصل على الشكل الانتشاري التالي للمتغيرين في نافذة المخرجات Output:

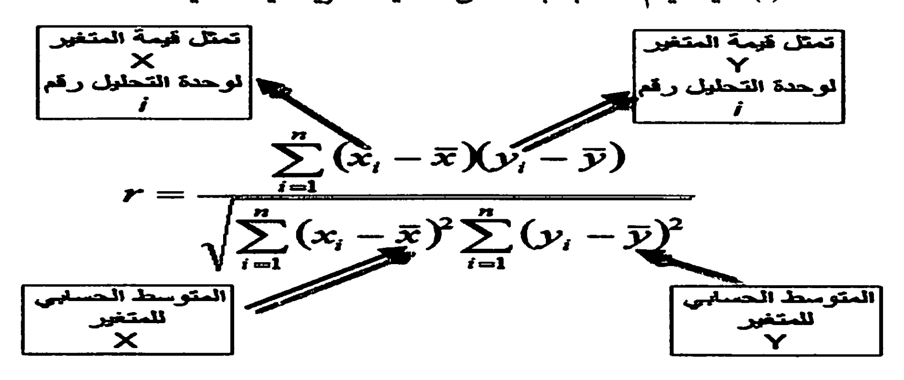
شكل رقم (٨-٥)

الشكل الانتشاري للمتغيرين درجة الطالب في القراءة (X) والمعدل التصحيلي (Y)



ويظهر لنا من الشـــكل الانتشاري أن هناك علاقة خطية موجبة (طردية) قوية بين المتغيريـــن. كما نلاحظ أنه لا توجد هناك قيم شـــاذة أو متطرفـــة خارجة أو منعزلة لوحدها بعيداً عن بقية البيانات. ومن ثم سنقوم بحساب الارتباط بين المتغيرين كمياً من خلال استخدام معامل ارتباط بيرسون.

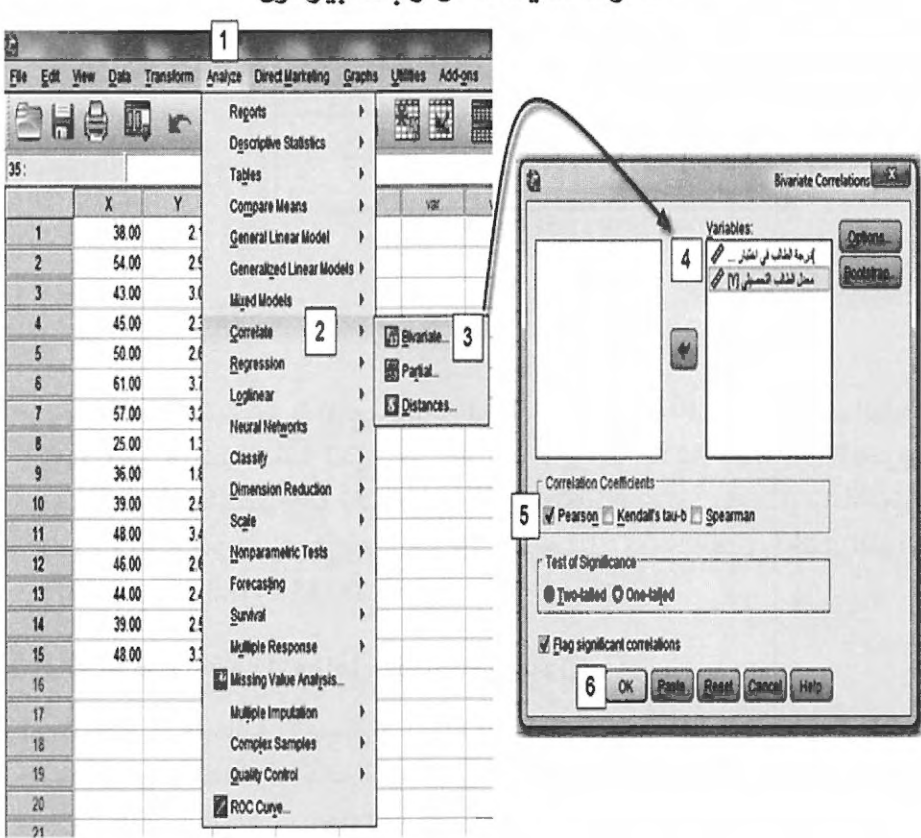
٢- اختيـــار معامل الارتباط المناســـب: وهنا ســـيتم اختيار معامل ارتباط بيرســـون Pearson (r) حيث يتم حسابه بناء على الصيغة الرياضية التالية:



$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

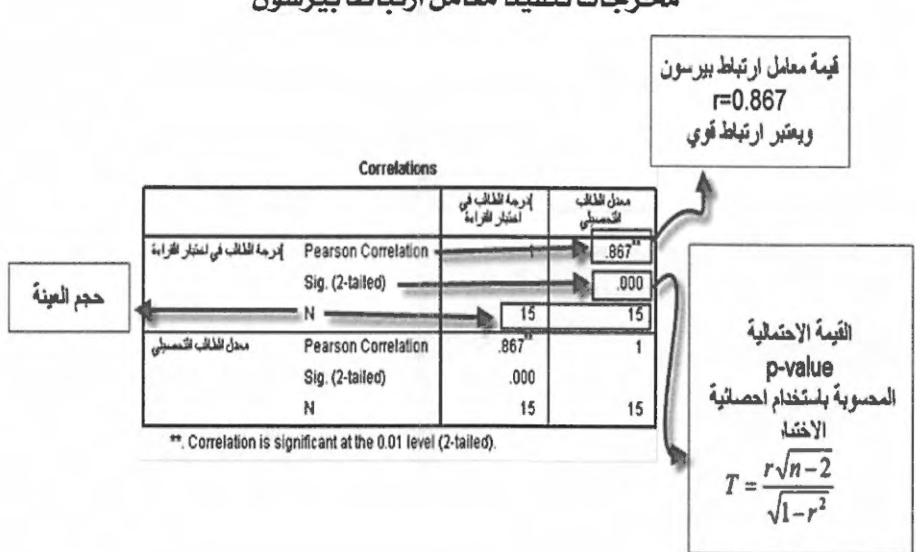
٣- الآن وبعــد إدخال المتغيــرات والبيانات إلى SPSS من خــلال نافذتي Variable View وData View كما في خطوة الحل رقم (١)، تتبع الخطوات حســب الأرقام التسلســلية مــن ١ إلى ٦ كما هو موضح بالشــكل (٨-٦) لتنفيــذ معامل ارتباط بيرسون Pearson.

شكل رقم (٨-٦)
خطوات تنفيذ معامل ارتباط بيرسون



وبعد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output كالتالي:

شكل رقم (٨-٧)
مخرجات تنفيذ معامل ارتباط بيرسون



Correlations

		درجة الطالب في اختبار القراءة	معدل الطالب التحصيلي
درجة الطالب في اختبار القراءة	Pearson Correlation	1	.867**
	Sig. (2-tailed)		.000
	N	15	15
معدل الطالب التحصيلي	Pearson Correlation	.867**	1
	Sig. (2-tailed)	.000	
	N	15	15

**. Correlation is significant at the 0.01 level (2-tailed).

٤- اتخاذ القرار وتفسير النتائج: بما أن P-value (0.01) > α (0.001) فإن الباحث سيرفض فرضية العدم التي تنص على أنه لا يوجد علاقة خطية بين المتغيرين. وهذا يقود إلى استنتاج أنه باستخدام معامل ارتباط بيرسون يوجد ارتباط معنوي قوي موجب بين متغيري معدل الطالب التحصيلي ودرجته في اختبار القراءة، $(r (13) = 0.87, p < 0.001)$.

إضاءات إحصائية حول معامل ارتباط بيرسون:

- لو نظرنا إلى إحصائية الاختبار T فإننا نلاحظ أن حجم العينة n يقع في البسط. وهذا يعني أنه كلما زاد حجم العينة زادت قيمة T بغض النظر عن قيمة r، وهذا بدوره يزيد بشكل أكبر احتمالية رفض فرضية العدم. أي أنه حتى وإن كانت قيمة r صغيرة جداً (ارتباط ضعيف جداً) فإنه وبزيادة حجم العينة يمكن الوصول إلى استنتاج أنه يوجد علاقة خطية معنوية إحصائياً بين متغيري الدراسة، ولكنها ليست معنوية عملياً. فمثلاً قد تكون قيمة $r = 0.10$ لمتغيرين ما، ومع ذلك يتم إحصائياً الوصول إلى استنتاج أنه يوجد ارتباط معنوي بين المتغيرين، وهذا غير مقبول منطقياً وعملياً. وهذا لا يعني أنه يتم تجاهل اختبار معنوية الارتباط بين

المتغيرات، فهو ضروري في حال أراد الباحث تعميم النتائج. ولكن عند إجراء اختبار معنوية الارتباط الخطي ينبغي توخي الحذر في تفسير النتائج والأخذ بعين الاعتبار كلاً من قيمة معامل الارتباط وحجم العينة.

- إذا تم تربيع قيمة r، فإننا نحصل على ما يسمى بمعامل التحديد r^2 والذي يفسر على أنه نسبة التباين أو التغير في أحد المتغيرين الذي يمكن تفسيره بالتغير في قيم المتغير الآخر ويعبر عنه عادة في صورة النسبة المئوية، ويستخدم كمؤشر على جودة أو دقة التنبؤ بقيم أحد المتغيرين (المتغير التابع) بالاعتماد على قيم المتغير الآخر (المتغير المستقل). وفي هذا التطبيق العملي فإن $r^2 =$ ٠,٧٥١٧ أي أن ٧٥,١٧٪ تقريباً من التباين أو التغير في متغير «المعدل التحصيلي للطالب» يفسر بـ / أو يعزى إلى التغير في متغير «درجة الطالب في القراءة».
- ينبغي رسم الشكل الانتشاري لمتغيري/ متغيرات الدراسة للتأكد من عدم وجود قيم شاذة والتي تؤثر على قيمة معامل الارتباط سواء بالزيادة أو النقص.
- تتأثر (تنقص) قيمة معامل ارتباط بيرسون بعدم تحقق شروط التوزيع الطبيعي، الخطية، وتجانس التباين التي سبق ذكرها.

تمرين تطبيقي (٨-١):

اختار باحث عينة عشوائية مكونة ٢٠ أسرة في أحد أحياء مدينة ما لدراسة العلاقة بين الدخل والإنفاق الشهري بآلاف الريالات، وكانت النتائج كما يلي:

جدول رقم (٨-٤)

توزيع (٢٠) أسرة في أحد أحياء مدينة ما حسب الدخل والإنفاق الشهري بآلاف الريالات

الدخل	١٠	١٥	١٨	١٣	١٢	١٥	١٢	١٣	٩	١١	٨	٩	١٠	١٣	١١	٩	٧	١٢	١٥	٨
الإنفاق	١٠	١٠	١٦	١١	١٠	١٥	٩	١٢	٥	٨	٨	٧	٩	١٠	٧	٩	٦	١١	١٠	٨

المطلوب:

١- وصف علاقة الارتباط بين متغيري الدخل والإنفاق الشهري باستخدام معامل ارتباط بيرسون r.

٢- حساب قيمة معامل التحديد R^2؟ تفسير معناه؟

٣- اختبرمعنوية الارتباط بين المتغيرين عند مستوى معنوية $\alpha =$ ٠,٠١.

تمرين تطبيقي (٨-٢):

في دراســة لبحث العلاقــة بين مفهوم الذات والطموح، اختيارات عينة عشــوائية مكونة من (٢١) فرداً، وكانت درجات المتغيرين لهؤلاء الأفراد كالتالي:

جدول رقم (٨-٥)
توزيع أفراد الدراسة حسب مفهوم الذات والطموح

مفهوم الذات	٢٢	٢٠	٢٠	١٩	١٨	١٧	١٨	١٨	١٧	١٦	١٦	١٥	١٦	١٤	٩	١٤	١٣	١٢	١١	١٧
الطموح	١٤	١٥	١٣	١٠	١٤	١١	١٢	١١	٨	١١	١٢	١١	٩	١٠	٥	٧	٨	٩	١٦	١٣

المطلوب:

١- وصف علاقة الارتباط بين متغيري مفهوم الذات والطموح باستخدام معامل ارتباط بيرسون r.

٢- حساب قيمة معامل التحديد R^2؟ تفسير معناه؟

٣- اختبار معنوية الارتباط بين المتغيرين عند مستوى معنوية $\alpha =$ ٠,٠١.

ثانياً - معامل الارتباط الجزئي بين المتغيرات Partial Correlation Coefficient:

مقدمة:

إن معامــل الارتباط الخطي البســيط بين المتغيريــن X وY لا يأخذ في الاعتبار احتمــال تأثير بعض المتغيرات الأخرى مثــل W، Z، ... إلخ على قوة وطبيعة العلاقة بــين المتغيريــن X وY. حيث يمكن أن نرى في حالات كثيــرة أن هناك علاقة خطية بين متغيرين، ولكن سرعان ما تختفي تلك العلاقة عند استبعاد أو ضبط تأثير متغير ثالــث أو أكثر. ومثال الطول والتحصيل العلمي للطالب الســابق ذكره خير مثال على ذلك. حيث وجد أن هناك علاقة خطية بين طول الطالب وتحصيله العلمي باستخدام معامل الارتباط الخطي البسيط والذي لم يأخذ بالاعتبار تأثير المتغيرات الأخرى مثل العمر، ولكن عند تثبيت أو اســتبعاد تأثير متغير العمر (أي حســاب معامل الارتباط بين متغيري التحصيــل والطول للطلاب عند عمر واحد محدد - مثلاً الطلاب الذين أعمارهم ١٥ عاماً) وجد أن ذلك الارتباط قد تلاشى. ولهذا السبب نجد أنه في كثير

من البحوث العلمية يفضل استخدام معامل الارتباط الجزئي (يمكن أن يتم ذلك أيضاً من خلال تحليل الانحدار الخطي المتعدد) على معامل الارتباط البسيط لاستبعاد أو ضبط التأثير المحتمل للمتغيرات الأخرى، ومن ثم يكون وصف علاقة الارتباط أدق وأصدق.

يرمز لمعامل الارتباط الجزئي بين المتغيرين X وY بعد استبعاد أو ضبط تأثير المتغير الثالث Z بالرمز $r_{xy.z}$، وهنا يسمى في هذه الحالة بمعامل الارتباط الجزئي من الرتبة الأولى لأنه تم ضبط متغير واحد فقط Z وصيغته الرياضية هي:

$$r_{xy.z} = \frac{r_{xy} - r_{xz} . r_{yz}}{\sqrt{(1-r_{xz}^2)(1-r_{yz}^2)}}$$

حيث r_{xy}, r_{xz}, r_{yz} ترمز لمعاملات ارتباط بيرسون من الرتبة صفر. وإذا تم ضبط متغيرين فإنه يسمى معامل ارتباط جزئي من الرتبة الثانية وهكذا. وعليه فإن معامل ارتباط بيرسون الخطي يسمى أحياناً بمعامل الارتباط الخطي البسيط، لأنه لم يأخذ في الاعتبار ضبط أو استبعاد تأثير أي متغيرات أخرى.

وهنا يمكن القول بأن الارتباط الجزئي يفيد في التالي:

- اكتشاف علاقات الارتباط الزائفة. أي اختفاء علاقة الارتباط بين متغيرين بعد استبعاد تأثير متغيرات أخرى.
- إبراز أو كشف الغطاء عن علاقات ارتباط خفية. أي الحصول على علاقة ارتباط لم يتم اكتشافها بمعامل الارتباط البسيط.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سيتم توضيح كيفية تنفيذ معامل ارتباط بيرسون الجزئي وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٨-٢):

البيانات في الجدول (٨-٦) تبين حجم الأصول والمبيعات والأرباح السنوية لإحدى السنوات بالدولار الأمريكي لمجموعة من الشركات العالمية الكبيرة.

جدول رقم (٨-٦)
حجم الأصول والمبيعات والأرباح بملايين الدولارات لـ (١٠) شركات عالمية

	الشركة									
	GM	Ford	Exxon	IBM	GE	Mobil	Phillip Morris	Chrysler	Du Pont	Texaco
الأصول	١٧٣٢٩٧	١٦٠٨٩٣	٨٣٢١٩	٧٧٧٣٤	١٢٨٣٤٤	٣٩٠٨٠	٣٨٥٢٨	٥١٠٣٨	٣٤٧١٥	٢٥٦٣٦
المبيعات	١٢٦٩٧٤	٩٦٩٣٣	٨٦٦٥٦	٦٣٤٣٨	٥٥٢٦٤	٥٠٩٧٦	٣٩٠٦٩	٣٦١٥٦	٣٥٢٠٩	٣٢٤١٦
الأرباح	٤٢٢٤	٣٨٣٥	٣٥١٠	٣٧٥٨	٣٩٣٩	١٩٠٩	٢٩٤٦	٣٥٩٢	٢٤٨٠	٢٤١٣

المصدر: http://www.uic.edu/classes/idsc/ids472/p1-4.xls .

والمطلوب:

١- إيجاد معامل ارتباط بيرسون البسيط بين متغيري الأصول والأرباح، ومن ثم اختبار معنويته الإحصائية عند $\alpha = ٠,٠٥$.

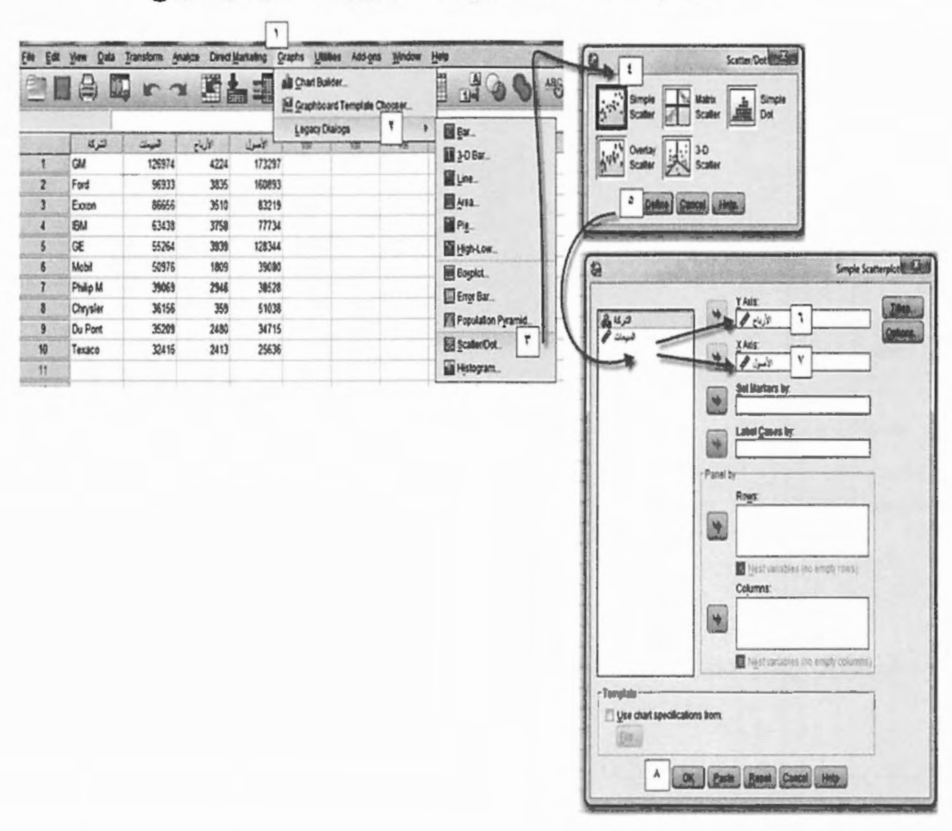
٢- إيجاد معامل ارتباط بيرسون الجزئي بين متغيري الأصول والأرباح بعد إزالة تأثير (ضبط تأثير) متغير «المبيعات»؟ ومن ثم اختبار معنويته الإحصائية عند $\alpha = ٠,٠٥$.

٣- مقارنة النتائج في ١ و٢.

خطوات الحل:

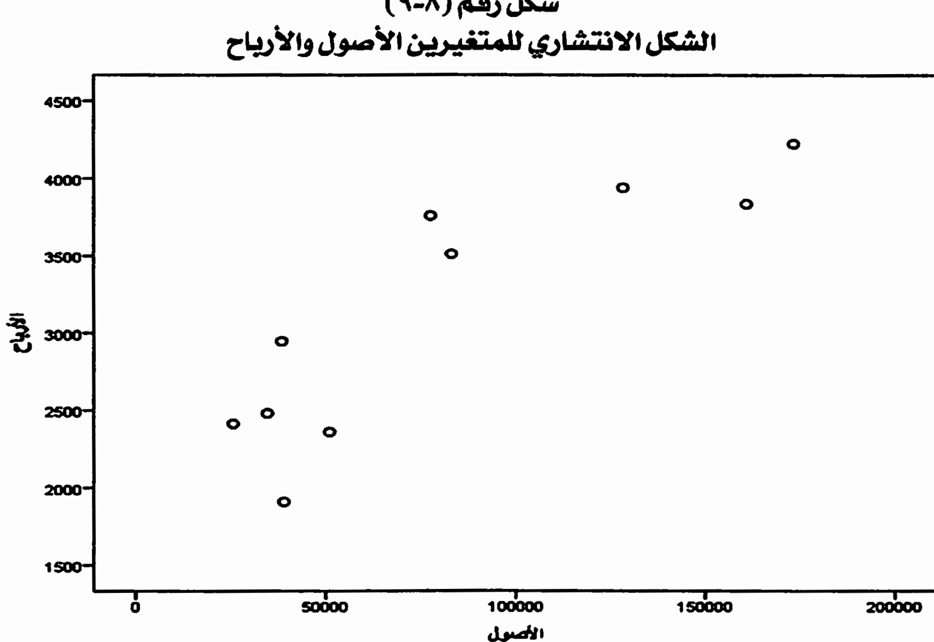
١- استخدم الشكل الانتشاري لتمثيل العلاقة بين متغيري الأصول والأرباح بيانياً. ولعمل ذلك قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٨ كما هو موضح بالشكل (٨-٨) للحصول على الشكل الانتشاري.

شكل رقم (٨-٨)
خطوات رسم الشكل الانتشاري للمتغيرين: الأصول والأرباح



وبعد التنفيذ نحصل على جدول الشـكل الانتشـاري التالـي للمتغيرين في نافذة المخرجات Output:

شكل رقم (٨-٩)
الشكل الانتشاري للمتغيرين الأصول والأرباح



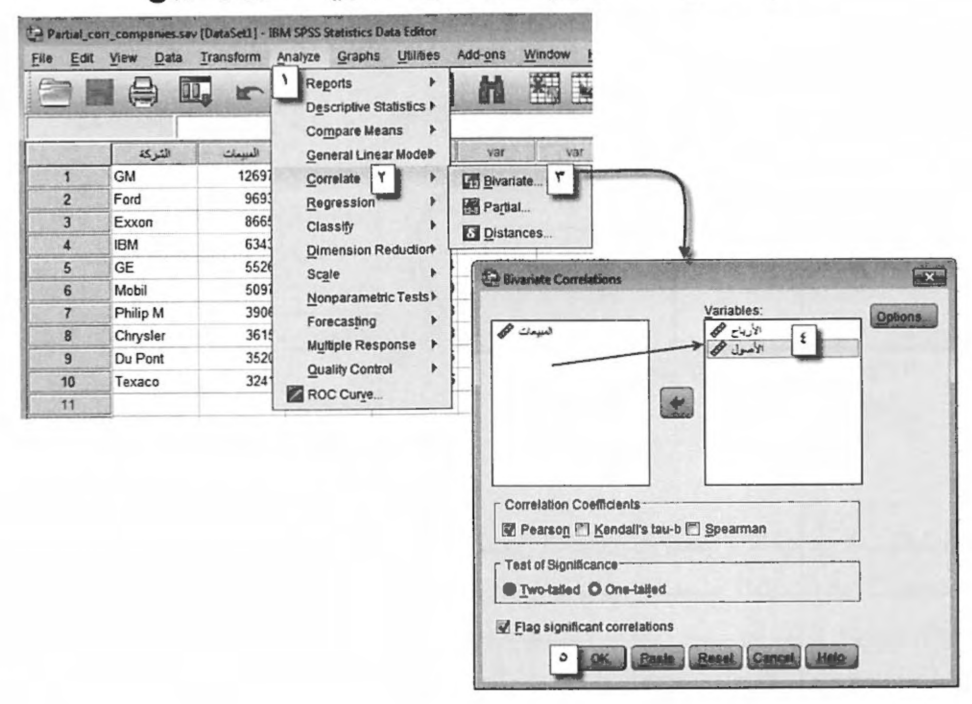
ويظهر لنا من الشـــكل الانتشاري أن هناك علاقة خطية موجبة (طردية) قوية بين المتغيرين.

٢- اختيـــار معامل الارتباط المناســـب: وهنا ســـيتم اختيار معامل ارتباط بيرســـون Pearson (r).

٣- الآن وبعـــد إدخال المتغيـــرات والبيانات إلى SPSS من خـــلال نافذتي Variable View وData View كما في خطوة الحل رقم (١)، تتبع الخطوات حســـب الأرقام التسلســـلية من ١ إلى ٨ كما هو موضح بالأشـــكال التاليـــة لتنفيذ معامل ارتباط بيرسون Pearson.

شكل رقم (٨-١٠)

خطوات تنفيذ معامل ارتباط بيرسون لمتغيري الأصول والأرباح



وبعد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output كالتالي:

شكل رقم (٨-١١)
مخرجات تنفيذ معامل ارتباط بيرسون لمتغيري الأصول والأرباح

Correlations

		الأرباح	الأصول
الأرباح	Pearson Correlation	1	.867**
	Sig. (2-tailed)		.001
	N	10	10
الأصول	Pearson Correlation	.867**	1
	Sig. (2-tailed)	.001	
	N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

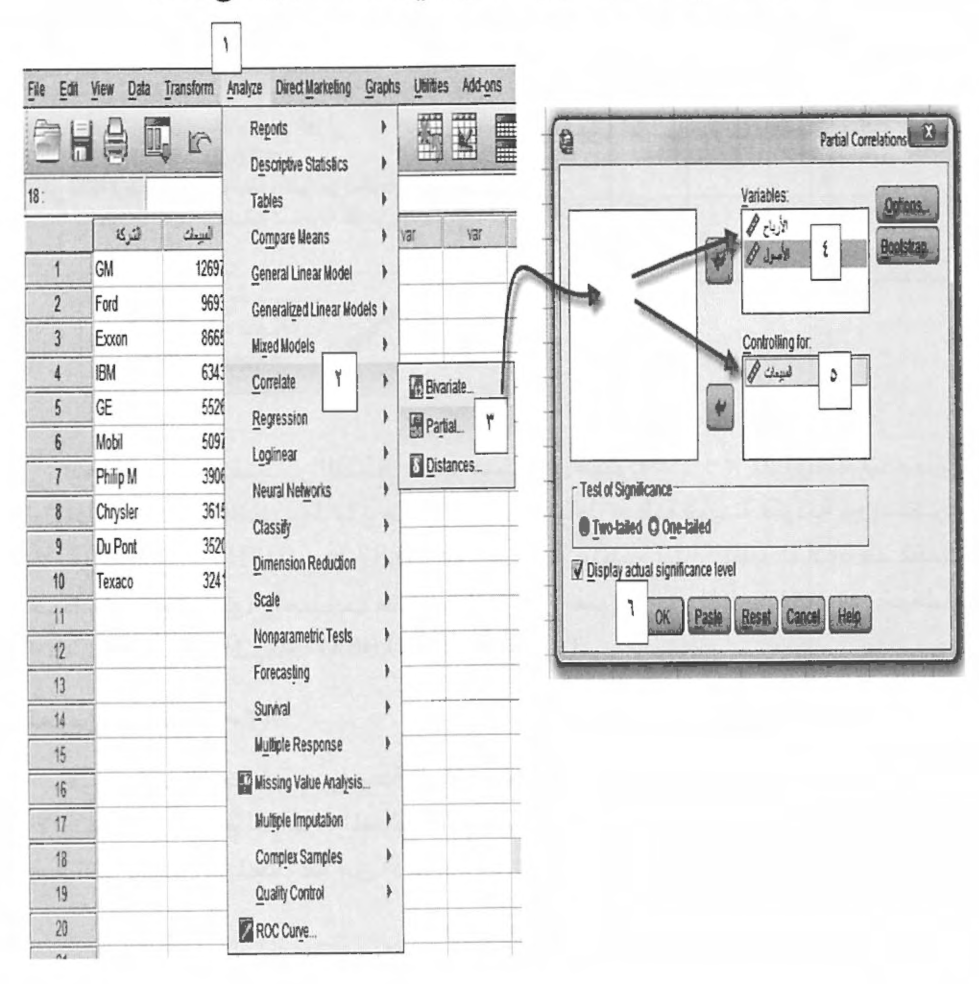
قيمة معامل ارتباط بيرسون
r=0.867
ويعتبر هذا ارتباط قوي طردي بين متغيري الأصول والأرباح

القيمة الإحتمالية المحسوبة من بيانات العينة
p-value=0.001

٤- الآن سـيتم تنفيذ معامل ارتباط بيرسـون الجزئي بين نفس المتغيرين السـابقين (الأصـول والأربـاح) ولكن بعد ضبط أو إزالة تأثير المتغيـر الثالث وهو المبيعات. حيـث إنه يعتبـر متغير المبيعات في الاقتصـاد متغيراً وسـيطاً يؤثر في العلاقة بين متغيري الأصول والأرباح. فليس بالضرورة أن تكون أرباح شـركة معينة ذات أصول أعلى أكبر من أرباح شركة أخرى ذات أصول أقل، وإنما هناك عوامل أخرى تؤثر على الأرباح بشـكل مباشـر أكثر من الأصول. ونلاحظ من المثال السابق أن معامل ارتباط بيرسـون البسـيط بين متغيري الأصول والأرباح $r = 0.867$ تشير إلـى وجود ارتباط خطي طردي قـوي بين المتغيرين. وهذه العلاقة قد تضعف أو تتلاشـى عندما نضبط أو نزيل تأثير المتغير الثالـث المتمثل في المبيعات. ولعمل ذلك إحصائياً فإننا سـنقوم بإيجاد معامل ارتباط بيرسـون الجزئي بين المتغيرين الأصول والأرباح مع ضبط تأثير المتغير الثالث - المبيعات.

ولتنفيذ ذلك باسـتخدام SPSS قم بتتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٦ كما هو موضح بالشكل (٨-١٢).

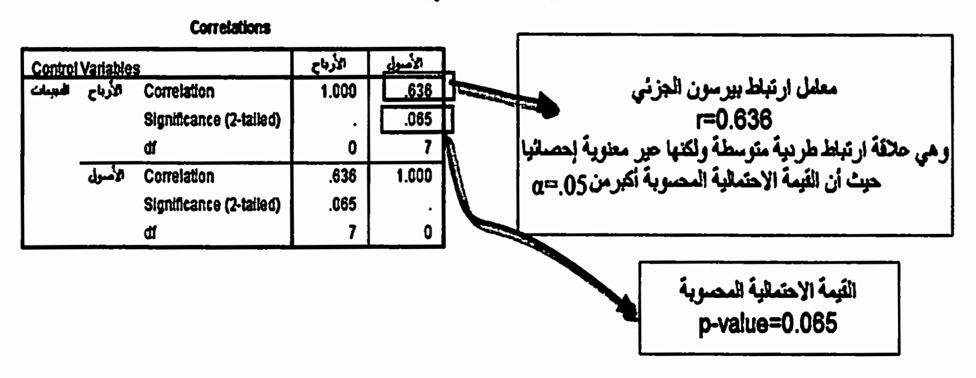
شكل رقم (٨-١٢)
خطوات تنفيذ معامل الارتباط الجزئي باستخدام برنامج SPSS



وبعد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output كالتالي:

شكل رقم (٨-١٣)

مخرجات تنفيذ معامل ارتباط بيرسون الجزئي بين المتغيرين الأصول والأرباح بعد ضبط تأثير متغير المبيعات



Correlations

Control Variables			الأرباح	الأصول
المبيعات	الأرباح	Correlation	1.000	.636
		Significance (2-tailed)	.	.065
		df	0	7
	الأصول	Correlation	.636	1.000
		Significance (2-tailed)	.065	.
		df	7	0

اتخاذ القرار وتفسير النتائج: عند مستوى معنوية٠,٠٥ = α، يوضح لنا معامل ارتباط بيرسون البسيط بأن هناك علاقة ارتباط خطية قوية طردية بين متغيري الأصول والأرباح $(r \approx 0.87, p = 0.001)$. في حين يتضح لنا أن تلك العلاقة قد فقدت جزءاً من درجة قوتها وأصبحت متوسطة وغير معنوية إحصائياً، وذلك عند ضبط أو إزالة تأثير متغير المبيعات $(r \approx 0.64, p = 0.065)$.

تمرين تطبيقي (٨-٣):

في أحد الأبحاث تم دراسة العلاقة بين المتغيرات الثلاثة التالية: المسئولية الاجتماعية، مفهوم الذات، والمستوى الاقتصادي. والدرجات في الجدول التالي تمثل قيم تلك المتغيرات لمجموعة من الأفراد.

جدول رقم (٨-٧)

بيانات مجموعة من الأفراد حسب متغيرات المسئولية الاجتماعية، مفهوم الذات، والمستوى الاقتصادي

	الأشخاص														
	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥
المسئولية الاجتماعية	١١	٦	١٥	٢٠	٤٠	١١	٣٠	١٦	٢٥	١٤	١١	١٥	٢٠	١٥	٢٥
مفهوم الذات	٧	٢	٢	٤	٨	٦	٧	٧	٥	٣	٨	٨	٩	٥	٤
المستوى الاقتصادي	٤	٦	٧	٧	٩	٩	٩	٦	٤	٥	٧	٨	٩	٧	٧

والمطلوب: حساب معامل الارتباط بين المتغيرين «المسئولية الاجتماعية» و«مفهوم الذات» بدون استبعاد (ضبط) تأثير المتغير الثالث «المستوى الاقتصادي»، ومقارنة ذلك بالعلاقة بين المتغيرين السابقين بعد استبعاد تأثير المتغير الثالث؟

مقاييس الارتباط اللامعلمية Non-Parametric Correlation Measures:

في حالة عدم ملاءمة استخدام أساليب الارتباط المعلمية لسبب أو أكثر مثل الالتواء الشديد في توزيع البيانات أو كون البيانات نوعية أو وجود القيم الشاذة، فإنه يتم استخدام البديل اللامعلمي للحصول على نتائج أكثر دقة ومصداقية. وفي هذا الجزء سيتم تقديم أكثر مقاييس الارتباط اللامعلمية البديلة وشرحها بطريقة عملية تفاعلية باستخدام برنامج SPSS.

أولاً - معامل ارتباط سبيرمان للرتب Spearman Rank Correlation Coefficient:

الهدف من استخدامه:

لوصف العلاقة الخطية بين متغيرين رتبيين، وذلك بتحديد مقدارها وقوتها واتجاهها. ويمكن أيضاً استخدامه مع المتغيرات الكمية.

ويعتمد معامل ارتباط سبيرمان على رتب البيانات وليس القيم الحقيقية لها في حسابه، حيث يتم ترتيب قيم المتغيرين تصاعدياً أو تنازلياً وفي حال تساوي القيم فإنه يتم إعطاء كل قيمة من تلك القيم متوسط رتب تلك القيم. ويرمز له بالرمز ρ وتنطق Rho أو الرمز r_s. وتقع قيمه بين -١ و+١ ويفسر بنفس الطريقة التي يفسر بها معامل ارتباط بيرسون.

وتجدر الإشارة إلى أنه يمكن اختبار معنوية الارتباط الخطي بين متغيري الدراسة باستخدام إحصائية الاختبار التالية:

$$T = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}}$$

وهي نفس إحصائية الاختبار المستخدمة لاختبار معنوية معامل ارتباط بيرسون مع استبدال r بـr_s حيث r_s يرمز لمعامل ارتباط سبيرمان، n حجم العينة، وT تتبع توزيع t بدرجات حرية n-2 (أي أن $T \sim t\ (n-2)$).

وإحصائية الاختبار هذه يتم استخدامها لاختبار فرضية العدم:

H_0: لا يوجد ارتباط خطي بين متغيري الدراسة (أي $H_0: \rho = 0$ حيث ρ ترمز لمعامل ارتباط بيرسون في حالة بيانات المجتمع وليس العينة).

مقابل فرضية البدل (الباحث):

H_1: يوجد ارتباط خطي معنوي بين المتغيرين ($H_1: \rho \neq 0$).

متى يستخدم:

يستخدم كبديل لا معلمي لمعامل الارتباط الخطي بيرسون r في حال كون البيانات رتبية أو في حالة مخالفة شـــروط التوزيع الطبيعي وتجانـــس التباين، وكذلك عندما تكون العلاقة بين المتغيرين غير خطية ولكنها رتيبة monotonic relationship.

شروط استخدامه:

١- أن تكون عينة البيانات عشوائية وبيانات كل متغير مستقلة بعضها عن بعض.

٢- أن تكـــون البيانات على هيئة أزواج مرتبـــة ordered pairs حيث لكل وحدة تحليل (متـــدرب أو طالب أو موظـــف مثلاً) يوجد قيمتان: (x,y) الأولى x تنتمي للمتغير الأول، والثانية y تنتمي للمتغير الثاني.

٣- أن تكون البيانات رتبية على الأقل. أي يمكن أن تكون البيانات فترية أو نسبية أيضاً.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي ســـيتم توضيـــح كيفية تنفيذ معامل ارتباط ســـبيرمان وتفســـير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٨-٣):

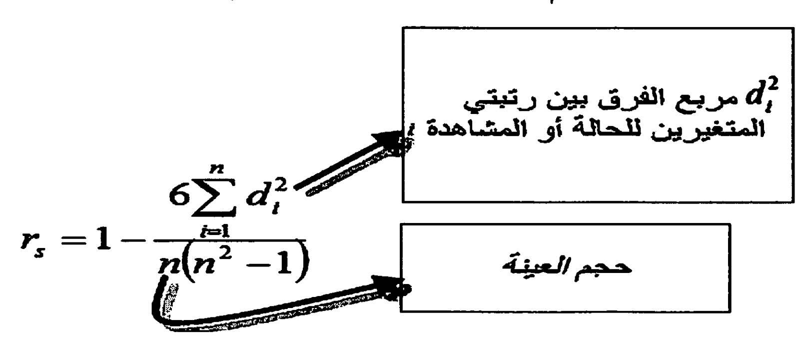
تمثل البيانات في الجدول (٨-٨) الوضع الاقتصادي لأســـرة الزوج وأسرة الزوجة (منخفض، متوســـط، جيد، جيد جداً، ممتاز) لعينة مكونة من (١٠) أسر تم اختيارهم عشوائياً (هندي، سلمان، ١٤٢٥هـ).

جدول رقم (٨-٨)
الوضع الاقتصادي لأسرة الزوج وأسرة الزوجة لعينة مكونة من ١٠ أسر

	الأسرة									
	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
الوضع الاقتصادي لأسرة الزوج	جيد	متوسط	جيد	ممتاز	جيدجداً	جيد	متوسط	ممتاز	جيد	جيد
الوضع الاقتصادي لأسرة الزوجة	جيد	منخفض	جيدجداً	جيدجداً	جيد	جيدجداً	متوسط	ممتاز	جيدجداً	منخفض

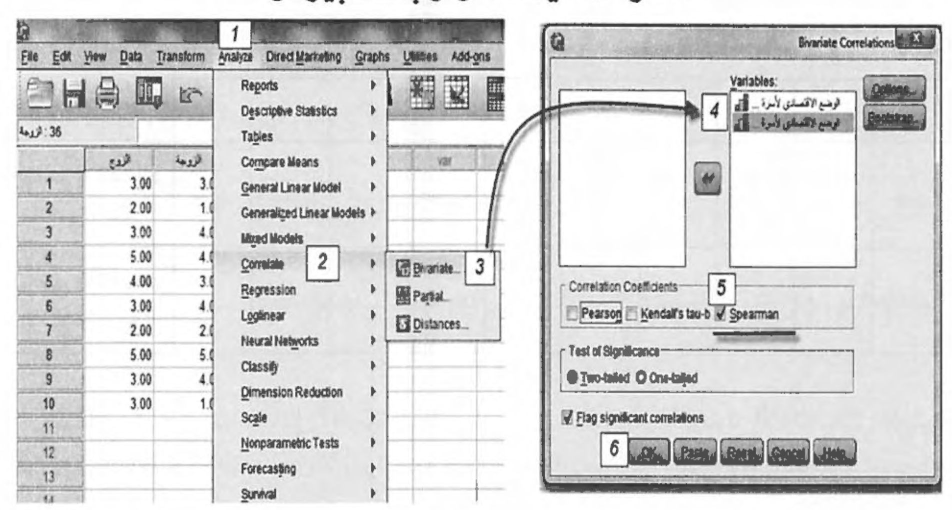
المطلوب: حساب معامل الارتباط الخطي بين المتغيرين «الوضع الاقتصادي لأسرة الزوج» و«الوضع الإقتصادي لأسرة الزوجة» واختبار معنوية هذا الارتباط عند مستوى معنوية $\alpha =$ ٠,٠٥؟

١- اختيار معامل الارتباط المناسب: وهنا سيتم اختيار معامل ارتباط سبيرمان للرتب Spearman (rs) حيث يتم حسابه بناء على الصيغة الرياضية التالية:



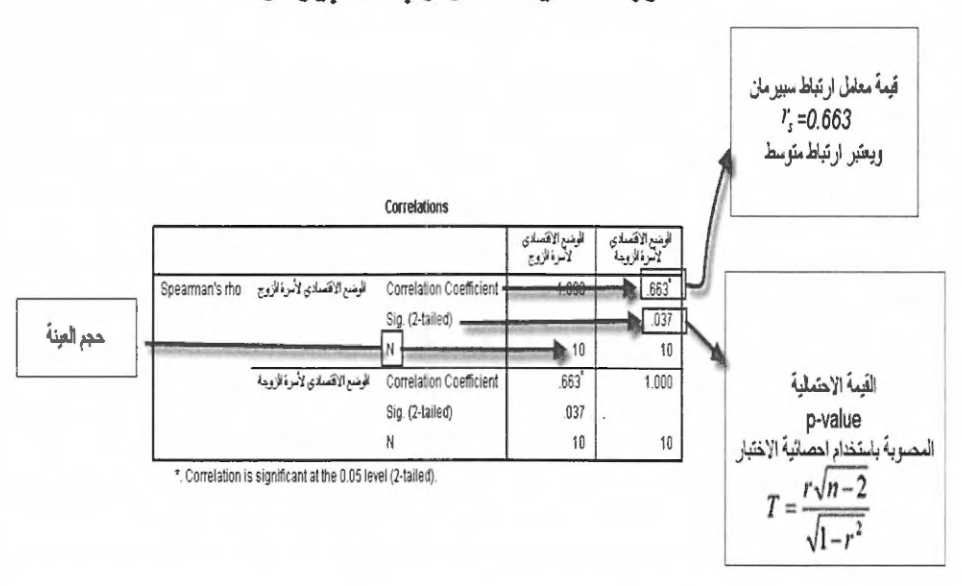
٢- قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View، ومن ثم قم بحفظ الملف، وهنا ينبغي أن يتم ترميز قيم المتغيرين رقمياً فيعطى ممتاز الرقم ٥، جيدجداً ٤، جيد ٣، متوسط ٢، ومنخفض١ وهنا سيتولى SPSS عملية إعطاء الرتب لقيم المتغيرين أثناء حساب قيمة معامل ارتباط سبيرمان. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٦ كما هو موضح بالشكل (٨-١٤) لتنفيذ معامل ارتباط سبيرمان Spearman (rs):

شكل رقم (٨-١٤)
خطوات تنفيذ معامل ارتباط سبيرمان



وبعد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output كالتالي:

شكل رقم (٨-١٥)
مخرجات تنفيذ معامل ارتباط سبيرمان



Correlations

			الوضع الاقتصادي لأسرة الزوج	الوضع الاقتصادي لأسرة الزوجة
Spearman's rho	الوضع الاقتصادي لأسرة الزوج	Correlation Coefficient	1.000	.663*
		Sig. (2-tailed)	.	.037
		N	10	10
	الوضع الاقتصادي لأسرة الزوجة	Correlation Coefficient	.663*	1.000
		Sig. (2-tailed)	.037	.
		N	10	10

*. Correlation is significant at the 0.05 level (2-tailed).

٣- اتخاذ القرار وتفسير النتائج: بما أن P-value (0.05) α < (0.037) فإن الباحث سيرفض فرضية العدم التي تنص على أنه لا يوجد ارتباط خطي بين المتغيرين. وهذا يقود إلى استنتاج أنه باستخدام معامل ارتباط سبيرمان يوجد ارتباط معنوي متوسط موجب بين متغيري «الوضع الاقتصادي لأسرة الزوج» و«الوضع الاقتصادي لأسرة الزوجة» $(r_s = 0.67, p = 0.037)$.

إضاءات إحصائية حول معامل ارتباط سبيرمان:

١- يعتبر معامل ارتباط سبيرمان حالة خاصة من معامل ارتباط بيرسون. والفرق بينهما أن معامل ارتباط سبيرمان يعتمد على رتب البيانات في حين يعتمد بيرسون على القيم الفعلية أو الحقيقية للبيانات.

- القيم الشاذة outliers ذات تأثير أقل على معامل سبيرمان مقارنة بمعامل بيرسون.
- يفضل استخدام معامل سبيرمان على معامل بيرسون للعينات الصغيرة الحجم.
- معامل كندال تاو - بي Kendall's tau-b هو معامل ارتباط آخر قدمه العالم كندال عام ١٩٣٨م لقياس علاقة الارتباط بين متغيرين رتبيين. في الغالب تكون قيمته العددية أقل من سبيرمان (conover, 1999). وقيمته تتراوح بين -١ و+١ ويفسر بنفس الطريقة التي يفسر بها معامل سبيرمان.

تمرين تطبيقي (٨-٤):

يحتوي الجدول (٨-٩) على نتائج التقييم وعدد الدورات التدريبية التي تم اجتيازها لعشرة من الموظفين العاملين بإحدى القطاعات الحكومية.

جدول رقم (٨-٩)
التقييم وعدد الدورات التي تم اجتيازها لعشرة من الموظفين

	الموظف									
	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
الدورات	٤	٣	٦	١	١	٣	٤	١	٢	٥
التقييم	مقبول	متوسط	ممتاز	مقبول	ضعيف	جيد	جيد	مقبول	مقبول	جيدجداً

المطلوب: دراســة الارتباط الخطي بين تقييم الموظف وعدد الدروات التي اجتازها باستخدام معامل ارتبط سبيرمان، ومن ثم اختبار معنوية الارتباط عند مستوى معنوية $\alpha = ٠,٠١$؟

تمرين تطبيقي (٨-٥):

يحتوي الجدول (٨-١٠) بيانات لـ (١٠) بنوك تجارية مرتبة حســب موجوداتها من العملة الأجنبية والتسهيلات الائتمانية التي تقدمها لعملائها.

جدول رقم (٨-١٠)
ترتيب ١٠ بنوك تجارية حسب موجوداتها من العملة الأجنبية والتسهيلات الائتمانية التي تقدمها لعملائها

	البنك									
	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
الترتيب حسب الموجودات	٢	٤	٥	٣	٦	١	٧	٨	١٠	٩
الترتيب حسب التسهيلات	١	٥	٦	٢	٤	٣	٨	٩	٧	١٠

المطلوب: حســاب الارتبــاط الخطي بين ترتيــب البنوك حســب موجوداتها من العملة الأجنبية والتسهيلات الائتمانية التي تقدمها لعملائها باستخدام معامل ارتبط سبيرمان، ومن ثم اختبار معنوية الارتباط عند مستوى معنوية $\alpha = ٠,٠١$؟

ثانياً - اختبار مربع كاي للاستقلالية Chi Square Test for Independence:

الهدف من استخدامه:

لاختبار الفرضية حول الارتبــاط المعنوي بين متغيرين نوعيين Two Categorical Variables.

شروط استخدامه:

١- أن تكون العينة مختارة بشكل عشوائي من مجتمع الدراسة.

٢- أن تكون المتغيرات محل الدراسة نوعية Categorical.

٣- أن تكون البيانات التي يتم تحليلها على شـكل تكـرارات ممثلة في جداول اقتران Contingency Table مكونة من صفوف وأعمدة.

٤- أن يتم تصنيف بيانات الوحدة المدروسـة (وحدة المعاينة) في خلية واحدة فقط من خلايا جدول الاقتران.

٥- أن لا يقل عدد الخلايا التي تكون تكراراتها المتوقعة أقل من خمسـة عن ٢٠٪ من مجموع الخلايا، وأن لا يكون هناك خلية تكرارها المتوقع أقل من واحد. في حال ظهور هذه المشكلة فإنه يمكن عمل التالي:

- دمـج الخلايا التي تكراراتها المتوقعة أقل من خمسـة مـع بعضها أو مع خلايا أخرى بشرط أن تكون الخلية الجديدة ذات معنى منطقي.
- أو استخدام طرق تحليل أخرى تتعامل مع هذه المشكلة.

إضاءة إحصائية حول مربع كاي للاستقلالية:

عند تحليل جداول الاقتران يلاحظ اسـتخدام اختبار مربع كاي للاسـتقلالية في حـال كون المتغيرين رتبيـين Ordinal أو كون أحدهما رتبياً Ordinal والآخر اسـمياً Nominal. هذا لا ينفي جدوى أو صحة النتائج إلى درجة معينة حين استخدام مربع كاي، ولكن يوجد هناك أدوات بديلة وتخصصية أكثر تكون نتائجها أكثر دقة ومصداقية لكونها تأخذ خاصية الترتيب للبيانات في الاعتبار، وهذا ما يفتقده مربع كاي.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سـيتم توضيح كيفية تنفيذ اختبار مربع كاي للاسـتقلالية بين المتغيرات النوعية وتفسـير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٨-٤):

أراد باحث في الشـؤون الأمنية دراسة ما إذا كان هناك علاقة أو ارتباط بين نوع الجريمة (سـرقة، خطف، قتل) والحالة الاجتماعية للأفراد (أعزب، متزوج، مطلق)، لذا قام بأخذ عينة عشـوائية من قائمة تضم أفراداً تم ضبطهم أمنياً لارتكابهم أحد الجرائم التالية: سـرقة (١٣٥فرداً)، خطـف (١٠٦ أفراد)، قتل (٤٩ فرداً). الجدول (٨-١١) يوضح التوزيع التكراري للأفراد حسب نوع الجريمة والحالة الاجتماعية:

جدول رقم (٨-١١)
توزيع عدد الأفراد المضبوطين أمنياً حسب نوع الجريمة والحالة الاجتماعية

		نوع الجريمة		
		سرقة	خطف	قتل
الحالة الاجتماعية	متزوج	١٠	٢٠	٣٠
	أعزب	٥	٨٠	١٥
	مطلق	١٢٠	٦	٤

خطوات الحل:

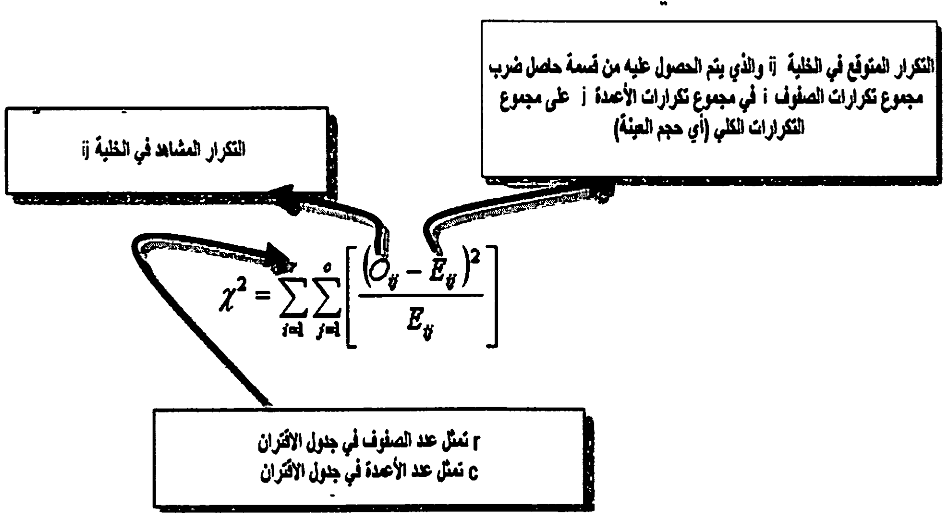
١- صياغة الفرضيات:

H_0: لا يوجد علاقة (ارتباط) معنوية بين نوع الجريمة والحالة الاجتماعية.

H_1: يوجد علاقة (ارتباط) معنوية بين نوع الجريمة والحالة الاجتماعية.

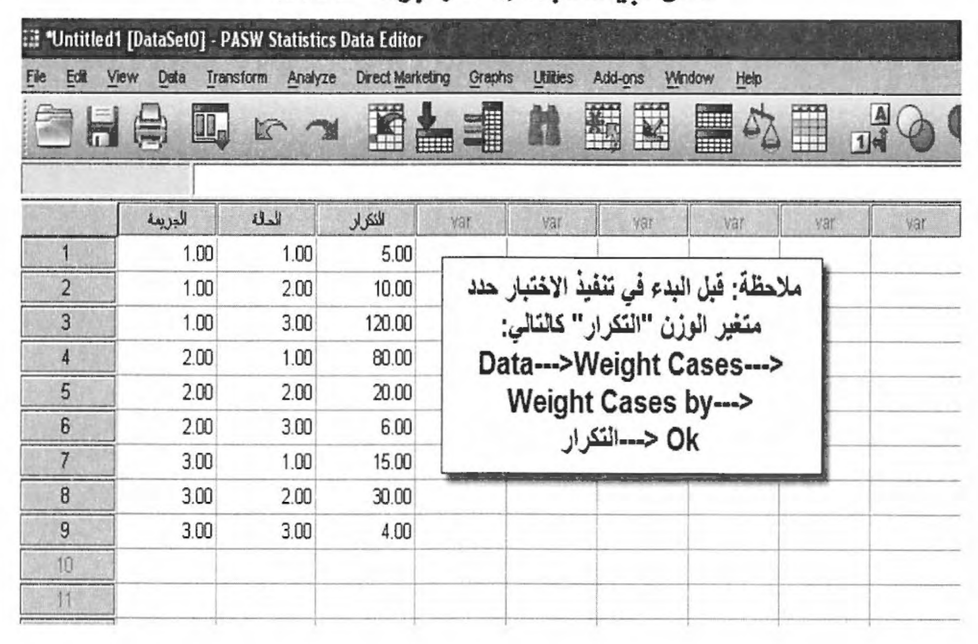
٢- تحديد مستوى المعنوية: وليكن $\alpha =$ ٠,٠٥.

٣- تحديد الاختبار الإحصائي المناسب: في هذه الحالة سيتم استخدام مربع كاي للاستقلالية Chi-Square Test for Independence. ومن ثم تأخذ إحصائية الاختبار الشكل التالي:



٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. والشكل (٨-١٦) يوضح ملف البيانات بعد إدخالها في SPSS:

شكل رقم (٨-١٦)
شكل البيانات بعد إدخالها بواسطة SPSS

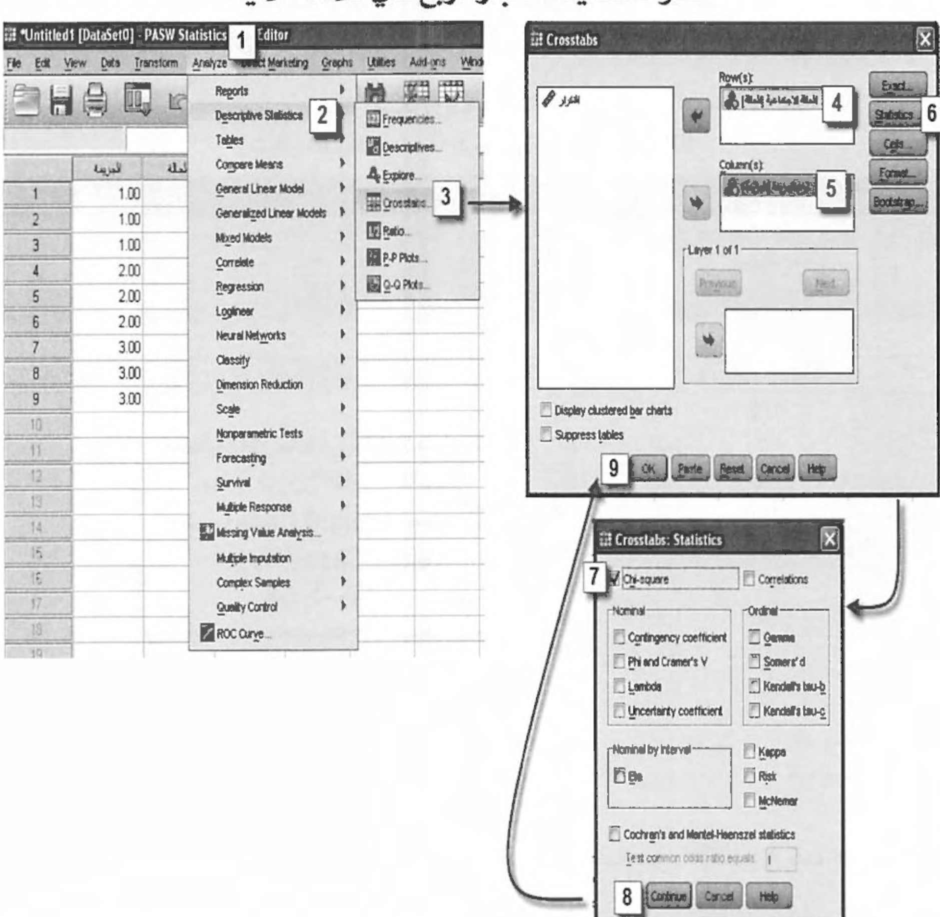


	الجريمة	الحالة	التكرار	var	var	var	var	var	var
1	1.00	1.00	5.00						
2	1.00	2.00	10.00						
3	1.00	3.00	120.00						
4	2.00	1.00	80.00						
5	2.00	2.00	20.00						
6	2.00	3.00	6.00						
7	3.00	1.00	15.00						
8	3.00	2.00	30.00						
9	3.00	3.00	4.00						
10									
11									

بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٨ كما هو موضح بالشكل (٨-١٧) لتنفيذ الاختبار:

شكل رقم (٨-١٧)

خطوات تنفيذ اختبار مربع كاي للاستقلالية



وبعد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output كالتالي:

شكل رقم (٨-١٨)
مخرجات تنفيذ اختبار مربع كاي للاستقلالية

➜ Crosstabs

[DataSet0]

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
الحالة الاجتماعية * نوع الجريمة	290	100.0%	0	.0%	290	100.0%

الحالة الاجتماعية * نوع الجريمة Crosstabulation

Count

		نوع الجريمة			Total
		سرقة	خطف	قتل	
الحالة الاجتماعية	أعزب	5	80	15	100
	متزوج	10	20	30	60
	مطلق	120	6	4	130
Total		135	106	49	290

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	248.977[a]	4	.000
Likelihood Ratio	266.147	4	.000
Linear-by-Linear Association	110.202	1	.000
N of Valid Cases	290		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.14.

٥- اتخاذ القرار وتفسير النتائج: بما أن α (0.05) < P-value (0.0001) فإن الباحث سيرفض فرضية العدم. وهذا يقود إلى استنتاج أنه عند مستوى معنوية $\alpha = ٠,٠٥$ فإن بيانات العينة تدل على وجود ارتباط معنوي بين متغيري نوع الجريمة والحالة الاجتماعية لمرتكب الجريمة $(\chi^2\ (4) = 248.977,\ p\text{-value} = 0.0001)$.

تمرين تطبيقي (٨-٦):

في إحدى المؤسسات الخاصة تم إجراء دراسة على ٤٠٠ موظف لدراسة العلاقة بين مستوى أدائهم في العمل (عالٍ، متوسط، متدنٍ) ومستوى نتائجهم في اختبار القدرات (عالٍ، متوسط، متدنٍ). والجدول (٨-١٢) يوضح توزيعهم حسب مستوى الأداء في العمل ونتائج اختبار القدرات.

جدول رقم (٨-١٢)
توزيع الموظفين حسب مستوى أدائهم في العمل ونتائجهم في اختبار القدرات

		مستوى الأداء في العمل		
		عالٍ	متوسط	متدنٍ
نتائج اختبار القدرات	عالٍ	٧٠	٨٠	٥٠
	متوسط	٣٠	٥٠	٤٠
	متدنٍ	٢٠	٥٠	١٠

المطلوب: هل هناك ارتباط معنوي بين مستوى الأداء في العمل ونتائج اختبار القدرات لموظفي تلك المؤسسة؟

تمرين تطبيقي (٨-٧):

في إحدى الدراسات أراد الباحث أن يتعرف على ما إذا كان هناك ارتباط بين متغيري «نوع السكن (ملك، إيجار) والمؤهل العلمي (فوق الجامعي، جامعي، ثانوي، أقل من ثانوي)». البيانات في الجدول (٨-١٣) تبين التوزيع التكراري لأفراد عينة عشوائية تم اختيارهم من مجتمع الدراسة.

جدول رقم (٨-١٣)
توزيع أفراد العينة حسب المؤهل العلمي ونوع السكن

		المؤهل العلمي			
		فوق الجامعي	جامعي	ثانوي	أقل من ثانوي
نوع السكن	ملك	٣	٩	٢	٢
	إيجار	١٠	١٦	١٥	١

المطلوب: عند مستوى معنوية $\alpha = ٠,٠٥$ وبناء على بيانات العينة هل هناك ارتباط معنوي بين المؤهل العلمي ونوع السكن؟

ملاحظة حول التمرين التطبيقي (٨-٧): عند التحليل ستلاحظ أن هناك أربع خلايا (٥٠٪ من مجموع الخلايا) لها تكرارات متوقعة أقل من (٥) وهذا بدوره ينافي أحد شروط استخدام اختبار مربع كاي للاستقلالية (أي يؤثر على مصداقية النتائج)، كيف يمكنك التغلب على هذا الخلل؟ يمكنك دمج الفئتين «ثانوي وأقل من ثانوي» ومن ثم تسميتها مثلاً «ثانوي فأقل» ودمج الفئتين «جامعي وفوق الجامعي» وتسميتها مثلاً «جامعي فأعلى»، ومن ثم ستلاحظ أن الخلايا التي لها تكرارات متوقعة أقل من خمسة قد تم معالجتها. أي أنك ستلاحظ أنه لن يكون لديك خلايا تكرارها المتوقع أقل من ٥.

ثالثاً - اختبار مربع كاي لجودة التوفيق Chi Square Goodness-of-Fit Test:

الهدف من استخدامه:

لاختبار الفرضية حول ما إذا كان توزيع البيانات في العينة متوافقاً مع توزيعها الافتراضي في المجتمع الذي سحبت منه العينة. بمعنى هل التوزيع التكراري المشاهد للبيانات Observed Frequency Distribution في العينة يتوافق مع التوزيع التكراري المتوقع Expected Frequency Distribution للبيانات في المجتمع الذي سحبت منه العينة. إذن يمكن القول بأن هذا الاختبار يستخدم إذا كان سؤال البحث يتعلق بتوزيع البيانات في المجتمع الدراسي.

شروط استخدامه:

١- أن تكون العينة مختارة بشكل عشوائي من مجتمع الدراسة.

٢- أن يكون المتغير محل الدراسة نوعياً Categorical (اسمي Nominal أو رتبي Ordinal) مع ملاحظة إمكانية استخدام مربع كاي للمتغيرات الكمية Quantitative (فتري أو نسبي Scale كما وردت تسميته في SPSS) ولكن بعد تحويل قيمها إلى فترات أو فئات.

٣- أن تكون البيانات التي يتم تحليلها على شكل تكرارات ممثلة في جدول اقتران Contingency Table مكون من صف واحد فقط مكون من خليتين أو أكثر.

٤- أن لا يتكرر ظهور بيانات الوحدة المدروسة (وحدة المعاينة) في أكثر من خلية واحدة من خلايا جدول الاقتران.

٥- أن لا يقل عدد الخلايا التي تكون تكراراتها المتوقعة أقل من خمسة عن ٢٠٪ من مجموع الخلايا. في حال ظهور هذه المشكلة فإنه يمكن للباحث دمج الخلايا التي تكراراتها المتوقعة أقل من خمسة مع بعضها أو مع خلايا أخرى بشرط أن تكون الخلية الجديدة ذات معنى منطقي أو يمكنه استخدام طرق تحليل أخرى تتعامل مع هذه المشكلة.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سيتم توضيح كيفية تنفيذ اختبار مربع كاي لجودة التوفيق وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٨-٥):

البيانات في الجدول التالي تمثل الحالة الاقتصادية لعينة مكونة من ٢٥ شخصاً تم اختيارها عشوائياً من مرتادي إحدى المجمعات التجارية، ويتوقع الباحث بأن توزيع الأفراد المرتادين لذلك المجمع التجاري حسب الحالة الاقتصادية متساوي، بمعنى أن نسبة (عدد) الأفراد الذين حالتهم الاقتصادية «ممتازة» تساوي نسبة (عدد) الأفراد الذين حالتهم الاقتصادية «جيدة» تساوي نسبة (عدد) الأفراد الذين حالتهم الاقتصادية «متوسطة» تساوي نسبة (عدد) الأفراد الذين حالتهم الاقتصادية «سيئة» تساوي ٢٥٪. هل تدعم بيانات العينة هذا التوقع أم لا؟ هل توزيع البيانات المشاهد في العينة يتوافق مع التوزيع المتوقع؟

جدول رقم (٨-١٤)
توزيع الأفراد حسب الحالة الاقتصادية

م	الحالة الاقتصادية	م	الحالة الاقتصادية	م	الحالة الاقتصادية	م	الحالة الاقتصادية	م	الحالة الاقتصادية
١	جيدة	٦	متوسطة	١١	متوسطة	١٦	ممتازة	٢١	جيدة
٢	متوسطة	٧	ممتازة	١٢	متوسطة	١٧	سيئة	٢٢	ممتازة
٣	ممتازة	٨	سيئة	١٣	ممتازة	١٨	جيدة	٢٣	متوسطة
٤	ممتازة	٩	ممتازة	١٤	ممتازة	١٩	جيدة	٢٤	متوسطة
٥	جيدة	١٠	سيئة	١٥	جيدة	٢٠	جيدة	٢٥	ممتازة

خطوات الحل:

١- صياغة الفرضيات:

H_0: لا يوجد اختلاف معنوي بين توزيع البيانات في العينة وتوزيعها الافتراضي في مجتمع الدراسة.

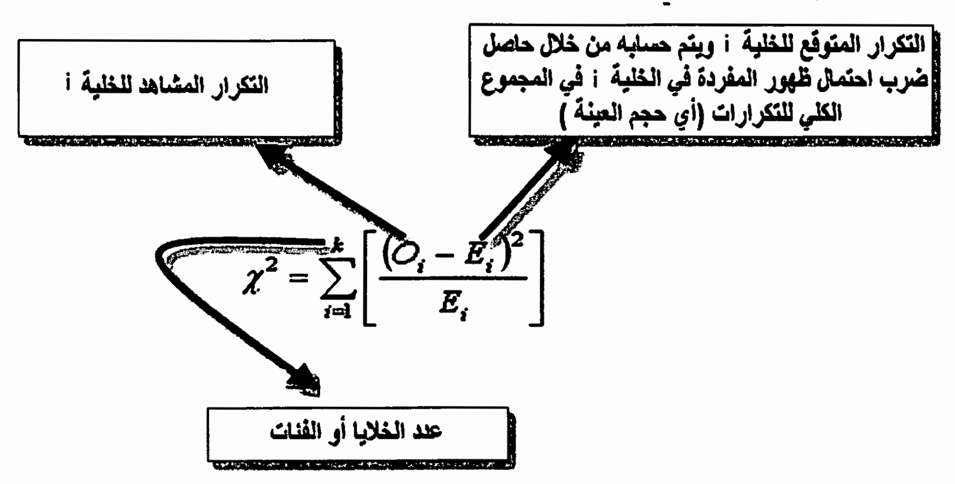
(بمعنى أنه يوجد تساوٍ في توزيع مرتادي ذلك المجمع التجاري حسب حالتهم الاقتصادية).

H_1: يوجد اختلاف معنوي بين توزيع البيانات في العينة وتوزيعها الافتراضي في مجتمع الدراسة.

(بمعنى أنه لا يوجد تساوٍ في توزيع مرتادي ذلك المجمع التجاري حسب حالتهم الاقتصادية).

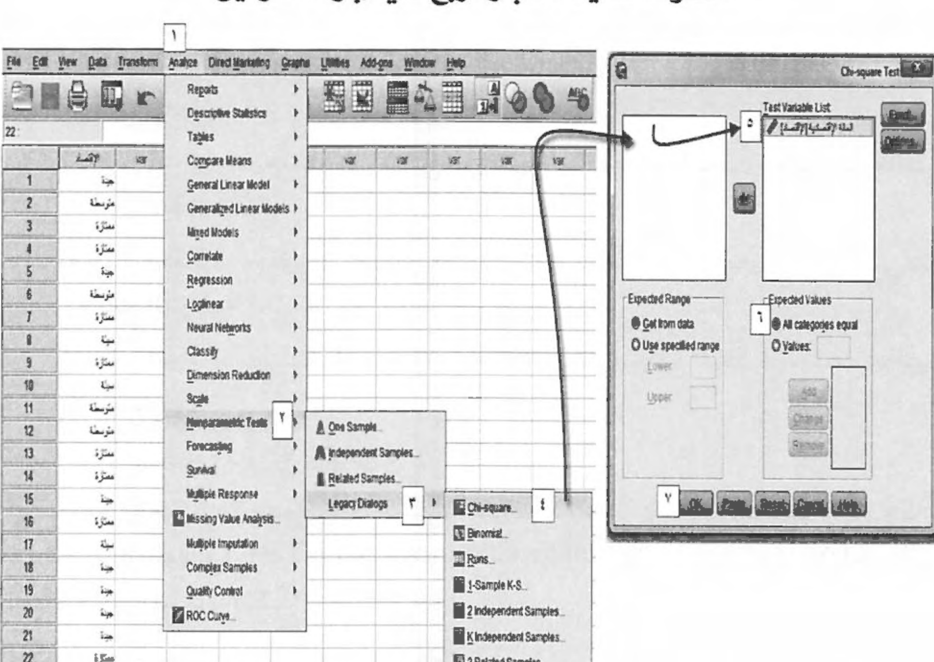
٢- تحديد مستوى المعنوية، وليكن $\alpha =$ ٠,٠٥.

٣- تحديد الاختبار الإحصائي المناسب: في هذه الحالة سيتم استخدام مربع كاي لجودة التوفيق Chi Square Goodness-of-Fit Test. ومن ثم تأخذ إحصائية الاختبار الشكل التالي:



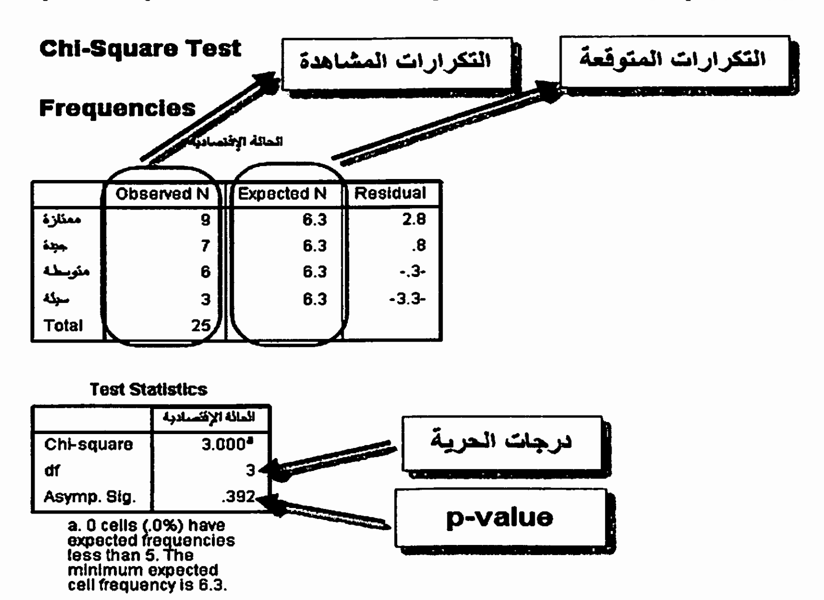
٤- الآن قم بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٧ كما هو موضح بالشكل (٨-١٩) لتنفيذ الاختبار.

شكل رقم (٨-١٩)
خطوات تنفيذ اختبار مربع كاي لجودة التوفيق



وبعد التنفيذ نحصل على جدول صغير في نافذة المخرجات Output كالتالي:

شكل رقم (٨-٢٠)
مخرجات تنفيذ اختبار مربع كاي لجودة التوفيق



	Observed N	Expected N	Residual
ممتازة	9	6.3	2.8
جيدة	7	6.3	.8
متوسطة	6	6.3	-.3-
سيئة	3	6.3	-3.3-
Total	25		

Test Statistics

	الحالة الإقتصادية
Chi-square	3.000[a]
df	3
Asymp. Sig.	.392

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 6.3.

٥- اتخاذ القرار وتفسير النتائج: بما أن $P\text{-value}\ (0.39) > \alpha\ (0.05)$ فإن الباحث لن يرفض فرضية العدم. وهذا يقود إلى استنتاج أنه عند مستوى معنوية $\alpha =$ ٠,٠٥ فإن بيانات العينة لا تدل على وجود اختلاف معنوي بين توزيع البيانات في العينة والمجتمع الدراسي الذي سحبت من العينة $(\chi^2\ (3) = 3, \text{p-value} = 0.39)$.

تمرين تطبيقي (٨-٨):

هناك اعتقاد بأن حوادث المركبات القاتلة تكون منتشرة بشكل أكبر في بعض أيام الأسبوع أكثر من غيرها. لذا قام باحث بأخذ عينة عشوائية من الأسابيع في إحدى السنوات في ولاية مونتانا الأمريكية من قاعدة بيانات معهد التأمين لسلامة الطرق السريعة Insurance Institute for Highway Saftey وقام بحصر الحوادث المصنفة على أنها حوادث قاتلة ومن ثم بدأ بحصر عدد الوفيات في كل حادث حسب اليوم في الأسبوع، وكانت النتائج كما في الجدول (٨-١٥) أدناه:

جدول رقم (٨-١٥)
توزيع عدد الوفيات حسب أيام الأسبوع

اليوم	السبت	الأحد	الإثنين	الثلاثاء	الأربعاء	الخميس	الجمعة
عدد الوفيات	٣٦	٣١	٢٠	٢٠	٢٢	٢٢	٢٩

المطلــوب: عند مســتوى معنوية $\alpha = ٠,٠٥$ هل تدعم هــذه البيانات الاعتقاد بأن حوادث المركبات المرورية القاتلة تتركز في أيام معينة من الأسبوع أكثر من غيرها؟

تمرين تطبيقي (٨-٩):

يمارس أحد المحلات نشــاطه التجاري ســتة أيام في الأسبوع ويعتقد مدير المحل أن الخدمــة الجيدة تجعل الزبائن يعودون مرة أخــرى، ولذلك دائماً يضع عدداً كافياً مــن الموظفين في الخدمة (١٥ موظفاً) كل يوم عمــل. كما أن لدى المدير اعتقاد بأن اليــوم (الســبت، الأحد، ....، الخميس) ليس له تأثير في حجــم العمل. ونظراً لتلقي مدير المحل بعض الشكاوي من الزبائن مفادها أن الخدمة في بعض الأيام تكون بطيئة مقارنة بالأيام الأخرى، فقد قام خلال الأسابيع الأربع والعشرين الماضية بجمع بيانات عن أعداد الزبائن خلال أيام الأسبوع كما هو موضح بالجدول (٨-١٦) أدناه (Triola, 2012, 547):

جدول رقم (٨-١٦)
توزيع عدد الزبائن المرتادين للمحل خلال الأربع والعشرين الأسبوع الماضية حسب أيام الأسبوع

	اليوم						
	السبت	**الأحد**	**الإثنين**	**الثلاثاء**	**الأربعاء**	**الخميس**	**المجموع**
عدد الزبائن	١٥٢٥	١٧١١	١٦٥٥	١٤٩٧	١٦٠٣	١٨٠١	٩٧٩٢

المطلوب: عند مســتوى معنوية $\alpha = ٠,٠١$ هل تدعــم هذه البيانات اعتقاد المدير بعدم تأثير اليوم في عدد الزبائن؟

الفصل التاسع
تحليل الانحدار الخطي Linear Regression Analysis

مقدمة:

يعــرف تحليل الانحدار على أنه أســلوب إحصائي لدراســة ونمذجة العلاقة بين المتغيــرات (Montgomery, Peck & Vining, 2006). حيث يتم دراســة انحدار أو اعتمــاد متغير واحد يســمى المتغير التابــع Dependent Variable على متغير واحد أو أكثــر تســمى بالمتغيرات المســتقلة Inependent Variables وبنــاء نموذج رياضي يصــف علاقة الارتباط بين المتغير التابع والمتغيرات المســتقلة. وقد ذكر مونتقومري Montgomery وآخرون أن أغراض تحليل الانحدار أربعة، وهي:

- وصــف البيانات Data Description. أي وصف شــكل علاقة الارتباط بين المتغير التابع والمتغيرات المستقلة.
- تقدير المعلمات Parameters Estimation. أي يتم استخدام تحليل الانحدار لتقدير معلمات بعض النماذج الرياضية المعروفة التي تربط بين متغير تابع ومتغيرات مستقلة معينة.
- التقديــر والتنبــؤ Prediction and Estimation. أي التقديــر والتنبؤ بقيم المتغير التابع بمعلومية قيم المتغيرات المستقلة.
- التحكــم Control. أي التحكم بقيم المتغير التابــع وفقاً لقيم محددة من المتغيرات المستقلة، وهذا الهدف الأخير يتم تحقيقه في حالة العلاقة السببية بين المتغير التابع والمتغيرات المستقلة.

إن المتصفح لأدبيات تحليل الانحدار يجد أنواعاً عديدة لتحليل الانحدار تناسب ظروفاً مختلفة متعلقة بنوع وعدد المتغيرات المستقلة والتابعة، وشكل وتوزيع المتغيرات وارتباطها ببعضها، وحجم العينة. وسنكتفي بالحديث في هذا الفصل على أكثر أساليب الانحدار اســتخداماً وأكثرها مألوفية لــدى كثير من الباحثين والدارســين وهو تحليل الانحدار الخطي. ينقسم الانحدار الخطي حسب عدد المتغيرات المستقلة إلى قسمين:

- الانحدار الخطي البسيط Simple Linear Regression حيث كلمة «الخطي» تعني أن العلاقــة خطية بين المتغيــر التابع ومعامل الانحدار، وكلمة «البســيط» تعني أن النموذج يحتوي على متغير مستقل واحد فقط.

- الانحدار الخطـــي المتعـــدد Multiple Linear Regression حيث «خطي» تعني أن معادلـــة الانحـــدار الخطي الموضحة أدنـــاه هي دالة خطية فـــي معاملات الانحدار و«متعدد» تعني أن تحليل الانحدار يحتوي على أكثر من متغير مستقل.

الصيغة الرياضية للانحدار الخطي المتعدد:

لنفـــرض أن لدينـــا عدد k من المتغيرات المســـتقلة $X1_i, X2_i, \ldots, Xk_i$ ومتغير تابع Y_i حيث $i = 1, 2, \ldots N$ دليل رقم المفردة أو الوحدة في مجتمع الدراســـة، فإن نموذج الانحدار الخطي المتعدد الاعتيادي يكتب كالتالي هو:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i$$

أما معادلة الانحدار الخطي المتعدد فهي:

$$E(Y_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}$$

حيث:

β_0 وβ_i (حيـــث β تقرأ بيتا) تمثل معامـــلات الانحدار الخطي، وهي معالم مجهولة ينبغي تقديرها من البيانات.

ε_i (حيث ε تقرأ إبسلون) يمثل حد الخطأ العشوائي وهو عبارة عن الفرق بين القيم المشاهدة Y_i والتركيبة الخطية للمتغيرات المستقلة $\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}$ أي أن $\varepsilon_i = Y_i - \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}$. وبعـــد تقديـــر معاملات الانحدار ومـــن ثم معادلة أو خط الانحدار من البيانات باســـتخدام أســـلوب المربعات الصغرى الاعتياديـــة Ordinary Least Squares أو أية طريقة أخرى كطريقة الإمكان الأعظم Maximum Likelihood يكتب نموذج الانحدار الخطي المقدر كالتالي:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \ldots + \hat{\beta}_k X_{ki} + \hat{\varepsilon}_i$$

ومن ثم تأخذ معادلة الانحدار الخطي المتعدد المقدرة الصيغة التالية:

$$\hat{E}(Y_i) = \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \ldots + \hat{\beta}_k X_{ki}$$

حيـــث العلامة "^" تقرأ «هات» ترمز للتقديـــر فمثلا $\hat{\beta}_k$ «حيث $\hat{\beta}$ تقرأ بيتا هات» ترمز لمقدر معامل الانحدار الخطي الجزئي β_k، و$\hat{\varepsilon}_i$ (وتكتب أيضا e_i) تســـمى بالبقايا Residulas وهي تعتبر تقديراً للخطأ العشوائي ε_i.

تفسير معاملات الانحدار الخطي:

- في حالة الانحدار الخطي البسيط $E(Y_i) = \beta_0 + \beta_1 X_i$ وكما هو موضح بالشكل (٩-١) فإن:

β_0 تعني الجزء المقطوع من المحور الرأسي (المتغير التابع) Y Intercept وهو عبارة عن متوسط المتغير التابع $E(Y)$ عندما تكون $x = 0$، ولكن إذا كان المتغير المستقل لا يحتوي بطبيعته على الصفر فإن تفسيره يصبح غير عملي.

β_1 يمثل ميل خط الانحدار على المحور الأفقي (المتغير المستقل) ويفسر على أنه مقدار التغير المتوقع في متوسط المتغير التابع الناتج من التغير بمقدار وحدة واحدة في المتغير المستقل.

- في حالة الانحدار الخطي المتعدد $E(Y)_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}$ فإن:

β_0 تعني الجزء المقطوع من المحور الرأسي Y وهي عبارة عن متوسط المتغير التابع $E(Y)$ عندما تكون قيم كل المتغيرات المستقلة تساوي الصفر.

$\beta_1, \beta_2, \ldots, \beta_k$ تمثل معاملات الانحدار الخطي الجزئية Partial Regression Coefficients. فمثلاً يفسر β_1 على أنه التغير المتوقع في متوسط المتغير التابع الناتج من التغير بمقدار وحدة واحدة في المتغير المستقل X1 مع ثبات التغير في المتغيرات المستقلة الأخرى في المعادلة. وتفسر بقية معاملات الانحدار الأخرى بنفس الطريقة.

اختبار المعنوية الإحصائية (اختبار الفرضيات) لمعادلة الانحدار الخطي المتعدد المقدرة:

هناك نوعان من الاختبارات حول معاملات الانحدار الخطي المتعدد هما:

١- اختبار المعنوية الإحصائية الكلية لمعادلة الانحدار الخطي المقدرة وتصاغ كالتالي:

الفرضية الصفرية (فرضية العدم) $H0: \beta_1, \beta_2, \ldots, \beta_k = 0$ وتعني أن جميع معاملات الانحدار الخطي الجزئية = الصفر (أي أن جميع المتغيرات المستقلة مجتمعة لا ترتبط بالمتغير التابع ارتباطاً معنوياً).

الفرضية البديلة $H1: \beta_i \neq 0$ حيث $i = 1, 2, \ldots, k$ أي أنه يوجد معامل انحدار خطي جزئي واحد على الأقل ≠ الصفر (أي أنه يوجد على الأقل متغير مستقل واحد من تلك المتغيرات المستقلة يرتبط بالمتغير التابع ارتباطاً معنوياً).

وهنا يتم اختبار H0 مقابل H1، حيث يتم مقارنة مستوى المعنوية الاسمي الذي يحدده الباحث سلفاً، وليكن $\alpha = 0.05$ مع القيمة المعنوية المحسوبة P-value في جدول تحليل التباين ANOVA والتي يرمزلها في مخرجات SPSS بـ Sig. لاتخاذ القرار الإحصائي حيال H0 بالرفض أو القبول (عدم الرفض).

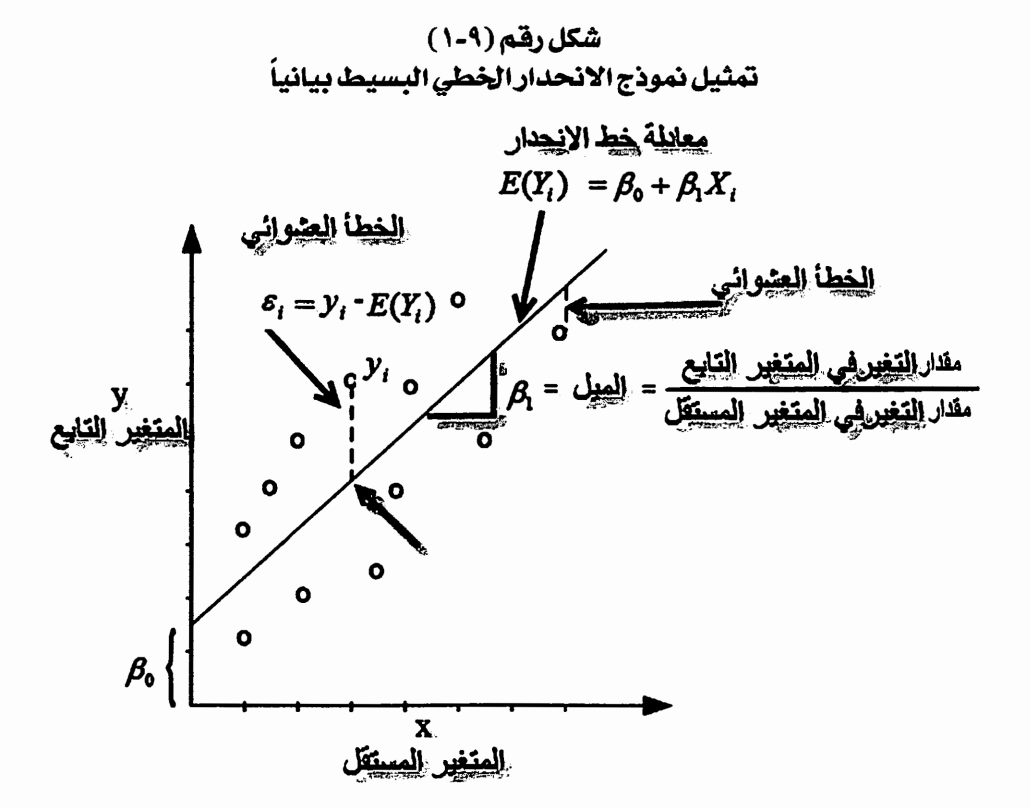
٢- اختبار المعنوية الجزئية لكل متغير مستقل على حدة. وهذا لا يتم إلا في حالة رفض فرضية العدم في اختبار المعنوية الكلية، حيث إنه عند رفض فرضية العدم وقبول الفرضية البديلة يتم البدء في البحث عن أي المتغيرات المستقلة مرتبط معنوياً بالمتغير التابع. ومن ثم تصاغ الفرضيات الإحصائية لكل متغير مستقل على النحو التالي:

الفرضية الصفرية (العدم) $H0: \beta_i = 0$ حيث $i = 1, 2, \dots, k$ وتعني أن معامل الانحدار الخطي الجزئي للمتغير المستقل X_i = الصفر (أي أن المتغير المستقل X_i لا يرتبط بالمتغير التابع ارتباطاً معنوياً).

الفرضية البديلة $H1: \beta_i \neq 0$ حيث $i = 1, 2, \dots, k$ وتعني أن معامل الانحدار الخطي الجزئي للمتغير المستقل X_i ≠ الصفر (أي أن المتغير المستقل X_i يرتبط بالمتغير التابع ارتباطاً معنوياً).

وبنفس الطريقة يتم اختبار H0 مقابل H1 حيث يتم مقارنة مستوى المعنوية الاسمي الذي يحدده الباحث سلفاً، وليكن $\alpha = 0.05$ مع القيمة المعنوية المحسوبة P-value في جدول معاملات الانحدار Coefficients والتي يرمز لها في مخرجات SPSS بـ Sig. لاتخاذ القرار الإحصائي حيال H0 بالرفض أو القبول.

شكل رقم (٩-١)
تمثيل نموذج الانحدار الخطي البسيط بيانياً



طرق إدخال المتغيرات المستقلة في نموذج الانحدار الخطي المتعدد:

هناك عدة طرق لإدخال المتغيرات المستقلة في نموذج الانحدار المتعدد. والاختيار من بين تلك الطرق يعتمد على خلفية الباحث النظرية أو خبرته العملية في علاقة المتغيرات المستقلة بالمتغير التابع، وكذلك عدد المتغيرات المستقلة. والجدير بالذكر أن طريقة وترتيب إدخال المتغيرات المستقلة تؤثر في نتائج التحليل وفي اختيار نموذج الانحدار الخطي النهائي، لذا ينبغي على الباحث الاهتمام باختيار الأسلوب الأنسب في دراسته. ويمكن تصنيف تلك الطرق إلى ثلاثة أقسام رئيسية (Field,2005; Ho, 2006) كالتالي:

أولاً - الطريقة التقليدية (طريقة الإدخال القصري) Standard or Forced Entry:

وهذه الطريقة تعتبر أكثر الأساليب استخداماً في الدراسات والبحوث حيث يتم فيها إدخال المتغيرات دفعة واحدة في نموذج الانحدار الخطي دون اعتبار ترتيب

إدخالها في النموذج، ومن ثم يتم تقييم درجة ارتباطها بالمتغير التابع وفق معايير معينة سيتم الحديث عنها لاحقاً. وهذه الطريقة يتم استخدامها عندما يكون لدى الباحث دراية مسبقة سواء من الدراسات السابقة أو الخلفية النظرية أو الخبرة العملية حول موضوع الدراسة. وهي تعتبر من أفضل طرق الانحدار لأنه يتم فيها تحديد النموذج أولاً ثم يتم جمع البيانات لتقدير معاملات الانحدار وبناء النموذج. وهذه الطريقة تناسب الحالات التي فيها عدد من المتغيرات المستقلة قليل.

ثانياً - الطريقة الهرمية Hierarchical Entry:

وفي هذه الطريقة يتم إدخال المتغيرات المستقلة في نموذج الانحدار الخطي فرادى أو في مجموعات بشكل متسلسل أو هرمي بناء على قوة ارتباطها أو أهميتها بالنسبة للمتغير التابع. وهذه الطريقة أيضاً تعتمد على خبرة الباحث وعلى الدراسات السابقة أو الخلفية النظرية لموضوع الدراسة.

ثالثاً - أسلوب الإدخال المتدرج Stepwise Entry:

ويتم استخدام هذه الطريقة في الدراسات الاستكشافية أو الاستطلاعية، وذلك عندما لا يكون لدى الباحث خلفية أو دراية كافية حول درجة ارتباط وأهمية المتغيرات المستقلة بالمتغير التابع. وهذه الطريقة تستخدم في مجال تنقيب البيانات Data Mining وفلترة المتغيرات وذلك عند توفر عدد كبير جداً من المتغيرات المستقلة ويريد الباحث تقليصها إلى العدد المناسب. وفي هذه الطريقة يتم إدخال المتغيرات المستقلة أو حذفها من النموذج بشكل متدرج وذلك بالاعتماد على معايير إحصائية بحتة. ومن عيوب هذه الطريقة أنها تعتمد بشكل كلي على بيانات العينة حيث إن نموذج الانحدار المبني بهذه الطريقة قد لا يصلح للاستخدام في حال تغيير العينة. ويتم هذا الأسلوب بعدة طرق وهي:

الاختيار للأمام Forward Selection:

وفي هذه الطريقة يتم البدء بنموذج انحدار يحتوي فقط على المعامل الثابت β_0 ثم يتم إدخال المتغير المستقل الأول والذي يعتبر الأكبر ارتباطاً معنوياً بالمتغير التابع، ثم يتم إدخال المتغير المستقل الثاني الأكبر ارتباطاً معنوياً بالمتغير التابع، وهكذا نستمر في الإضافة بشكل متدرج حتى يصبح ارتباط المتغير المستقل الحالي غير معنوي. وفي حالة دخول المتغير المستقل في النموذج فإنه في هذه الطريقة يبقى في النموذج ولا يتم حذفه.

طريقة الحذف للخلف Backward Deletion:

وهذه الطريقة عكس الطريقة السابقة حيث يتم إدخال جميع المتغيرات المستقلة مبدئياً في النموذج ثم نبدأ بحذف المتغير المستقل الأقل ارتباطاً بالمتغير التابع بشرط أن يكون ارتباطه بالمتغير التابع غير معنوي ثم نقوم بحذف المتغير المستقل الثاني الأقل ارتباطاً وغير معنوي بالمتغير التابع، وهكذا نستمر في الحذف بشكل متدرج حتى نصل إلى المرحلة التي تكون فيها المتغيرات المستقلة المتبقية في النموذج مرتبطة بالمتغير التابع ارتباطاً معنوياً. وفي حالة حذف المتغير المستقل فإنه في هذه الطريقة لا يتم إعادته مرة أخرى إلى النموذج.

الطريقة المتدرجة Stepwise:

وهي تجمع بين الطريقتين السابقتين حيث يتم البدء كما في حالة الاختيار للأمام بنموذج انحدار يحتوي المعامل الثابت β_0 فقط، ثم يتم إدخال المتغير المستقل الأكبر ارتباطاً معنوياً بالمتغير التابع، وفي الخطوة التالية يتم ادخال المتغير المستقل الثاني الأكبر ارتباطاً معنوياً بالمتغير التابع، وهنا يتم اختبار معنوية ارتباط المتغيرين الأول والثاني ويتم حذف المتغير غير المرتبط معنوياً بالمتغير التابع، وهكذا نستمر في الإضافة والحذف حتى نصل إلى نموذج الانحدار الذي يحتوي فقط المتغيرات المستقلة المرتبطة به معنوياً.

ويلاحظ أنه ليس بالضرورة أن تعطي تلك الطرق نفس النموذج النهائي للانحدار الخطي، والسبب في ذلك أن ترتيب إدخال المتغيرات في النموذج يلعب دوراً كبيراً في تحديد درجة الارتباط الجزئي للمتغير المستقل بالمتغير التابع في ظل وجود المتغيرات المستقلة الأخرى الموجودة في النموذج.

الهدف من استخدام الانحدار الخطي المتعدد:

في الواقع العملي ولا سيما في العلوم الاجتماعية والإنسانية يستخدم الانحدار الخطي المتعدد للأهداف التالية:

١- لإيجاد أفضل معادلة انحدار للتنبؤ بقيم المتغير التابع بمعلومية قيم المتغيرات المستقلة.

٢- تحديد المتغيرات المستقلة المرتبطة معنوياً بالمتغير التابع ومدى مقدار المساهمة الجزئية والكلية للمتغيرات المستقلة في تفسير التباين في المتغير التابع.

٣- التقدير الكمي للارتباط المتعدد بين المتغير التابع والمتغيرات المستقلة مجتمعة.

متى يستخدم:

عندما يكون الهدف بناء نموذج رياضي (معادلة الانحدار الخطي) لوصف وتفسير العلاقة بين متغير تابع كمي من المستوى الفتري أو النسبي ومجموعة من المتغيرات المستقلة الكمية أو النوعية، أو كليهما معاً واستخدام ذلك النموذج في التنبؤ بقيم المتغير التابع بمعلومية قيم المتغيرات المستقلة.

شروط استخدامه:

إن بناء معادلة الانحدار الخطي المتعدد تتطلب مجموعة من الشروط حيث الشروط من ١ إلى ٤ ضرورية لصلاحية المعادلة في تعميم نتائج العينة على مجتمع الدراسة (Field, 2005; Montgomery, Peck & Vining, 2006; Miles & Shevlin, 2007).

١- أن تكون العلاقة بين المتغير التابع والمتغيرات المستقلة خطية.

٢- أن يكون توزيع الخطأ العشوائي طبيعياً لكل قيمة من قيم المتغير التابع (أو لكل مجموعة من قيم المتغيرات المستقلة).

٣- أن يكون متوسط الخطأ العشوائي صفراً، وتباينه ثابتاً أو متجانساً لكل قيمة من قيم المتغير التابع (أو لكل مجموعة من قيم المتغيرات المستقلة).

٤- أن تكون قيم الخطأ العشوائي مستقلة بعضها عن بعض.

وهناك شروط أخرى يجب تحققها لكونها ضرورية لعملية جودة التوفيق للمعادلة ولدقة التقدير والتنبؤ بقيم المتغير التابع، ومنها:

٥- عدم وجود مشكلة الازدواج الخطي (الارتباط الخطي المشترك المتعدد) Multicollinearity بين المتغيرات المستقلة حيث يقصد بالازدواج الخطي الارتباط الخطي الكبير بين متغيرين مستقلين أو أكثر.

٦- خلو البيانات من القيم الشاذة Outliers والمؤثرة Influential values والقيم الرافعة Leverage values التي تؤثر على دقة نموذج الانحدار الخطي.

قبل البدء في تحليل الانحدار الخطي يجب معالجة البيانات وتهيئتها للتحليل، حيث يجب التأكد من خلوها من أخطاء تسجيل البيانات وأخطاء إدخالها في الحاسب، ومعالجة القيم المفقودة، كما ينبغي أن يكون حجم العينة مناسباً، وقد سبق الحديث عنه في الفصل الخامس.

ويتم استخدام البقايا التي يرمز لها بالرمز e والتي تمثل الفرق بين القيم المشاهدة أو الفعلية للمتغير التابع $\hat{y}$ والقيم التقديرية أو التنبؤية لها $\hat{y}$ باستخدام المعادلة التقديرية Fitted equation (أي $e = y - \hat{y}$) لدراسة مدى تحقق تلك الافتراضات.

والجدول (٩-١) التالي يوضح شروط تحليل الانحدار الخطي وطرق فحصها وتقييمها ومعالجتها. كما أن الشكل (٩-٢) يوضح لنا شكل انتشار البقايا مقابل القيم التنبؤية في حالة تحقق الشروط، بالإضافة إلى بعض الأشكال الممكن ظهورها في الواقع العملي، وذلك في حالة مخالفة بعض تلك الشروط.

جدول رقم (٩-١)

شروط تحليل الانحدار وطرق فحصها وتقييمها ومعالجتها

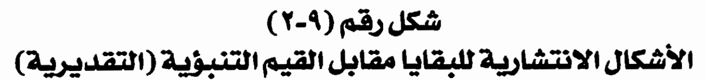
الشرط Assumption	وسائل التقييم وطرق العلاج الممكنة
١- الخطية Linearity أي العلاقة الخطية بين المتغير التابع والمتغيرات المستقلة.	الرسم البياني للبقايا Residual Plots وذلك من خلال الشكل الانتشاري للبقايا والقيم المقدرة $\hat{y}$. كما يمكن استخدام الشكل الانتشاري لدراسة العلاقة بين كل متغير مستقل والمتغير التابع. وربما تكون القيم الشاذة أو المؤثرة هي من أحد أسباب عدم الخطية. وفي حالة عدم تحقق هذا الشرط فإنه ينبغي معالجة الوضع إما بإضافة بعض المتغيرات مثل تربيع المتغير المستقل المسبب للمشكلة أو ضرب متغير مستقل بآخر، أو باستخدام التحويلات الخطية Linear Transformation على المتغيرات المستقلة أو التابع أو كليهما حسب الحالة، أو استخدام نماذج انحدار غير خطية، أو أساليب الانحدار اللامعلمية.
٢- التوزيع الطبيعي للخطأ العشوائي Normality of Random Errors.	الرسم البياني للبقايا وذلك من خلال المدرج التكراري للبقايا، والرسم أو الشكل الاعتدالي أو الطبيعي للبقايا Normality Plot أو الرسم البياني q-q. كذلك يتم بطرق غير بيانية (اختبارات احصائية) لاختبار الاعتدالية للخطأ العشوائي منها على سبيل المثال اختبار كولمقورف - سميرنوف Kolmogorov-Smirnov واختبار شابيرو - ويلك Shapiro-Wlik. وتجدر الإشارة إلى أن تأثير عدم تحقق هذا الشرط طفيف لا يكاد يذكر إذا كان الانحراف عن الاعتدالية غير شديد وكان حجم العينة كبير نسبياً. ويمكن معالجة عدم الاعتدالية باستخدام التحويلات الخطية المناسبة.
٣- تجانس تباين الخطأ العشوائي عند كل قيمة من قيم المتغير التابع Homogeneity of Random Errors.	الرسم البياني للبقايا وذلك من خلال الشكل الانتشاري للبقايا والقيم المقدرة $\hat{y}$، كذلك استخدام اختبار ليفين لتساوي التباينات Levene's Test for Equality of Variances. والسبب الأكثر احتمالية لاختلاف التباين هو أن يكون التباين دالة في متوسط المتغير التابع فقد يزداد أو يتناقص تباين الخطأ العشوائي بتغير قيم متوسط المتغير التابع. ومن الآثار المترتبة على مخالفة هذا الشرط هو أن معاملات الانحدار المقدرة بأسلوب المربعات الصغرى تقل كفاءتها، مما يؤدي إلى تضخم في الخطأ المعياري لها أكثر من اللازم، وهذا يضعف احتمالية الكشف عن المتغيرات المستقلة المرتبطة معنوياً بالمتغير التابع. وفي حالة عدم تحقق هذا الشرط فإنه ينبغي معالجة الوضع إما باستخدام التحويلات الخطية على المتغير المستقل أو التابع - وفي الواقع العملي قد يؤدي إجراء التحويل الخطي على المتغير التابع إلى معالجة المشكلة - أوباستخدام تحليل الانحدار باستخدام المربعات الصغرى الموزونة Weighted Least Squares Regression Analysis أو استخدام نماذج انحدار غير معلمية.

تابع - جدول رقم (٩-١).

الشرط Assumption	وسائل التقييم وطرق العلاج الممكنة
٤- استقلالية قيم الخطأ العشوائي عن بعضها البعض Independence of Random errors.	إن استقلالية (عدم ارتباط) قيم الخطأ العشوائي (او قيم المتغير التابع) عن بعضها البعض يتأثر بشكل كبير بتصميم المعاينة المستخدم في الدراسة. ففي حالة المعاينة العشوائية البسيطة تكون احتمالية استقلال القيم عن بعضها كبيرة بينما في تصاميم المعاينة التي يتم فيها اختيار مفردات الدراسة على شكل مجموعات Clusters تكون قيم الخطأ العشوائي أو قيم المتغير التابع عرضة للارتباط، لذا ينبغي الحذر في مثل هذه الحالة وفحص استقلالية البيانات بعناية. ومن طرق تشخيص استقلالية الخطأ العشوائي: - الاستعلام عن نوعية المعاينة المستخدمة. أيضا الاستعلام حول ما إذا كانت البيانات تتبع ترتيب زمني أو مكاني معين. فالمعاينة التي يتم فيها اختيار مفردات العينة على شكل مجموعات أو تتبع فيها البيانات تسلسل زمني أو مكاني في الظهور تكون مدعاة للشك والريبة حول استقلالية البيانات. عدا ذلك يمكن بوجه عام القبول باستقلاليتها. - الرسم البياني للبقايا وذلك من خلال الشكل الانتشاري للبقايا والقيم المقدرة لا المرتبة حسب ظهورها الزمني أو المكاني، وفي هذه الحالة يمكن استخدام اختبار دوربن - واتسن Durbin-Watson. وعند عدم تحقق شرط الاستقلالية فإنه ينبغي استخدام أساليب إحصائية أخرى مثل نماذج السلاسل الزمنية.
٥- عدم وجود مشكلة الازدواج الخطي Multicollinearity بين المتغيرات المستقلة.	- فحص مصفوفة الارتباط Correlation Matrix أي فحص الارتباط الخطي بين كل متغيرين مستقلين. ويتفق المختصون تقريباً على أن مشكلة الازدواج الخطي تعتبر حاضرة إذا كانت قيمة معامل الارتباط الخطي بين متغيرين مستقلين أكبر أو تساوي ٠.٧٠. وينبغي أن يكون فحص مشكلة الازدواج الخطي من أولى الخطوات في تحليل الانحدار، لأنه يؤدي إلى مشاكل عدة منها: عدم دقة المعلومات التي نحصل عليها من الرسومات البيانية لطبيعة العلاقة بين المتغيرات المستقلة والتابع، تضخم تباين معاملات الانحدار، وهذا يجعل المتغيرات المستقلة المعنوية تبدو غير معنوية الارتباط بالمتغير التابع، كما أنه ربما يؤدي إلى تغيير اتجاه العلاقة بين المتغير المستقل والمتغير التابع بالشكل الخاطئ، ومن ثم فإن هذا يؤدي إلى عدم دقة نموذج الانحدار المقدر. - استخدام مؤشر «معامل تضخم التباين VIF». فإذا كانت قيمته أكبر من ١٠ فإن ذلك ربما يعود لوجود مشكلة الازدواج الخطي. - مقارنة القيمة الاحتمالية الكلية المحسوبة (Global p-value) للنموذج مع القيمة الاحتمالية المحسوبة الجزئية للمتغيرات المستقلة منفردة. فإذا كانت القيمة الاحتمالية الكلية للنموذج معنوية إحصائياً والمتغيرات المستقلة منفردة غير معنوية إحصائياً، فإن ذلك ربما يعود لوجود مشكلة الازدواج الخطي بين متغيرين أو أكثر من المتغيرات المستقلة. ومن الحلول المقترحة لمعالجة تلك المشكلة: - جمع بيانات إضافية من خلال زيادة حجم العينة. - حذف بعض أو دمج المتغيرات العالية الارتباط مع مراعاة عدم الإخلال بالدراسة أو النموذج. - تقليص عدد متغيرات الدراسة بطريقة تعالج مشكلة الارتباط الخطي العالي بينها باستخدام أساليب إحصائية متقدمة مثل التحليل العاملي Factor Analysis أو تحليل المكونات الأساسية Principal Component Analysis. - استخدام أسلوب تحليل انحدار آخر يسمى «تحليل التل» Ridge Regression.

تابع - جدول رقم (٩-١).

الشرط Assumption	وسائل التقييم وطرق العلاج الممكنة
٦- خلو النموذج من القيم الشاذة Outliers والمؤثرة Influential.	- استخدام الرسم البياني للبقايا وذلك من خلال الشكل الانتشاري للبقايا والقيم المقدرة $\hat{Y}$. - فحص البقايا والبقايا المعيارية Studentized or Standardized Residuals وبوجه عام يفضل استخدام البقايا المعيارية من النوع Studentized. وتختلف طريقة ودرجة تأثير القيم الشاذة على دقة التقدير في النموذج فمثلاً تسمى القيمة الشاذة بالقيمة الرافعة Leverage إذا كانت القيمة الشاذة تقع على امتداد خط الانحدار بعيداً عن باقي القيم كما في الشكل ٩-٣ (أ). وهنا يكون تأثير القيمة الشاذة على تقدير معاملات الانحدار ضيئلاً جداً، ولكنها قد توحي لنا بأنه يوجد ارتباط خطي بين المتغير المستقل والتابع مع عدم وجود هذه العلاقة في الحقيقة كما يوضحه الشكل (٩-٣ «ب»). لذا فإن الاعتماد فقط على المؤشرات الكمية لوصف علاقة الارتباط بين المتغيرات قد يكون مضلل مع وجود القيم الرافعة لذا لا بد من اكتمال الصورة لوصف علاقة الارتباط بين المتغيرات باستخدام الأشكال الانتشارية. أما إذا كانت القيمة الشاذة بعيدة عن بقية القيم ولا تقع على امتداد خط الانحدار فإنها تسمى في هذه الحالة بالقيمة المؤثرة Influential ومن ثم فإن لها تأثيراً سلبياً كبيراً على معاملات الانحدار. - المؤشر الإحصائي Cook's D: ويتم حسابه لكل مفردة من مفردات العينة فإذا كانت $D > 1$ (أكبر من الواحد الصحيح) فإن تلك المفردة تعتبر قيمة مؤثرة. - المؤشر الإحصائي Standardized DfFit (الفرق المعياري في القيمة التنبؤية): وفيه يتم حساب الفرق في القيمة التنبؤية لكل مفردة أو حالة بعد استبعادها من النموذج فإذا كان $\lvert S\tan dardized\ Dfit \rvert > 2\sqrt{p/n}$، حيث p عدد معاملات الانحدار في النموذج وn حجم العينة، فإن المفردة مؤثرة. وذكر (Field, 2005) أن القيمة الشاذة تعتبر مؤثرة إذا كانت القيمة المطلقة للفرق المعياري في القيمة التنبؤية أكبر من الواحد الصحيح أي إذا كان $\lvert S\tan dardized\ Dfit \rvert > 1$. - المؤشر الإحصائي Standardized DfBeta (الفرق المعياري في معامل الانحدار): وفيه يتم حساب الفرق المعياري في قيمة معامل الانحدار الناتج من استبعاد مفردة معينة. فإذا كانت $\lvert S\tan dardized\ DfBeta \rvert > 2/\sqrt{n}$ فإن الحالة مؤثرة. - واقترح (Field, 2005) أن القيمة الشاذة تعتبر مؤثرة إذا كانت القيمة المطلقة للفرق المعياري في معامل الانحدار أكبر من الواحد الصحيح أي إذا كان $\lvert S\tan dardized\ DfBeta \rvert > 1$ كما اقترح Stevens (1992) أن القيمة الشاذة تعتبر مؤثرة إذا كانت $\lvert S\tan dardized\ DfBeta \rvert > 2$. - ومن أساليب معالجة هذه المشكلة استخدام التحويلات الخطية، أو استخدام أساليب انحدار أخرى لا تتأثر بالقيم الشاذة مثل الانحدار المانع Robust Regression. - وبوجه عام فإنه يمكن وبسهولة التأكد من درجة تأثير القيم الشاذة على نموذج الانحدار من خلال توفيق النموذج في وجود القيم الشاذة، ومن ثم توثيق النموذج بعد استبعاد القيم الشاذة ومن ثم مقارنة النتائج مع وبدون القيم الشاذة.



شكل رقم (٩-٢)

الأشكال الانتشارية للبقايا مقابل القيم التنبؤية (التقديرية)

e البقايا

0

القيم التنبؤية $\hat{y}$

(أ)

e

0

$\hat{y}$

(ب)

e

0

$\hat{y}$

(ج)

e

0

$\hat{y}$

(د)

e

0

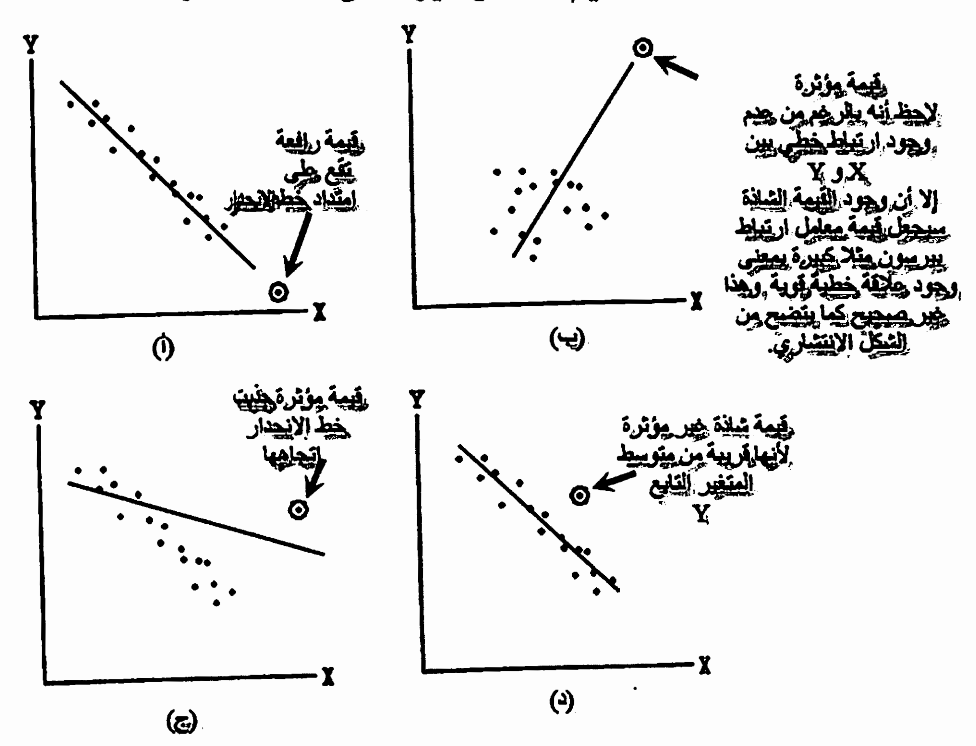
$\hat{y}$

(هـ)

حيث: (أ) يمثل النمط العشوائي Random Pattern في حالة تحقق الشروط، (ب) و(ج) شرط العلاقة الخطية غير متحقق، (د) شرط تجانس التباين غير متحقق، (هـ) شرط الاستقلالية غير متحقق.

شكل رقم (٩-٣)

أنماط مختلفة للقيم الشاذة وتأثيرها على خط الانحدار



إضاءات إحصائية حول الانحدار الخطي:

١- ينبغي أن ندرك أنه في الواقع العملي لا تتحقق تلك الشـــروط بالشكل التام، وذلك لطبيعـــة تعقيدات التطبيقات العملية. ولكـــن يعتبر نموذج تحليل الانحدار المقدر مفيداً إذا لم يتم مخالفة تلك الشروط بشكل كبير وفي ظل وجود الحجم المناسب للعينة (Agresti, Finlay, 2009).

٢- في حالة مخالفة أحد أو بعض شروط الانحدار الخطي فإن من وسائل المعالجة التي يتم اللجوء لها أولاً هو إجراء التحويلات الخطية للإبقاء على استخدام الانحدار الخطي الاعتيادي، وذلك لسهولة التعامل معه وسهولة تفسيره ومألوفيته لدى الباحثين ومحللي البيانات، ولكن بشرط أن تكون التحويلات الخطية المستخدمة غير معقدة مما يجعل عملية تفسير النتائج لاحقاً صعبة، وربما غير مفهومة أو مضللة لدى الكثير من الباحثين. وبوجه عام فإنه يتم استخدام التحويلات الخطية على المتغير المستقل لتصحيح مشكلة «عدم الخطية». في حين تستخدم التحويلات الخطية على المتغير التابع لتصحيح مشكلة «عدم تجانس التباين» و«عدم الطبيعية أو الاعتدالية»، وهذا ربما يؤدي إلى المساعدة في تحسين درجة الخطية. أيضاً استخدام التحويلات الخطية على المتغير التابع لتصحيح مشكلة «عدم الطبيعية» ربما يساعد في تصحيح مشكلة «عدم تجانس التباين». ومن التحويلات الخطية الأكثر استخداماً لتحويل المتغير التابع أو المتغير/ المتغيرات المستقلة ما يلي:

- التحويل اللوغاريتمي $Log\ (Y)$ أو $Log\ (X)$.
- الجذر التربيعي $\sqrt{Y}$ أو $\sqrt{X}$.
- تربيع أو حساب مقلوب قيم المتغير المستقل X^2 أو $\frac{1}{X}$.
- هناك أساليب إحصائية متقدمة لإجراء التحويلات الخطية تسمى بتحويلات القوى Power Transformations ومنها تحويل Box-Cox Transformation التي تستخدم لإجراء التحويلات الخطية على المتغير التابع Y، وتحويل Box-Tidwell والذي يستخدم لإجراء التحويلات الخطية على المتغيرات المستقلة. ويمكن تنفيذ هذه التحويلات باستخدام برمجيات خاصة Macros باستخدام إحدى برامج التحليل الإحصائي الشهيرة مثل SAS وSPSS أو R. والمستخدم العادي غير المتخصص يصعب عليه إجراؤها بدون مساعدة واستشارة إحصائي متخصص.

٣- ينبغي توخي الحذر عند استخدام معادلة الانحدار الخطي المقدرة في التنبؤ أو تقدير قيم المتغير التابع بناء على قيم المتغيرات المستقلة التي تقع بعيداً نسبياً خارج مدى بيانات المتغيرات المستقلة المستخدمة في بناء أو توفيق المعادلة، حيث يمكن أن تتغير طبيعة وشكل العلاقة بين المتغير التابع والمتغيرات المستقلة حسب مدى بيانات المتغيرات المستقلة. فإذا كان الباحث يتوقع أن العلاقة الخطية للانحدار مستمرة خارج نطاق بيانات دراسته أو بحثه، فإنه يمكنه الاطمئنان لاستخدام معادلة الانحدار الخطي في التنبؤ.

الخطوات العملية لتحليل الانحدار الخطي المتعدد:

١- بناء النموذج Model Building.

٢- التحقق من كفاية أو ملاءمة النموذج Model Adequacy.

٣- تشخيص وتقييم النموذج Model Diagnostics، أي التحقق من عدم مخالفته للشروط السابق ذكرها (١-٦).

٤- صلاحية النموذج Model Validation لاستخدامه في التنبؤ. ومن الأساليب المتبعة في التأكد من صلاحية النموذج المقدر ما يلي:

- فحص وتحليل معاملات الانحدار ومقارنتها مع خبرات الباحث و/أو الخلفية النظرية لموضوع البحث.
- جمع بيانات جديدة وتطبيق النموذج عليها لاختبار قدرته التنبؤية.
- تجزئ البيانات التي سبق جمعها ومن ثم استخدام جزء منها لبناء نموذج الانحدار الخطي واختبار قدرته التنبؤية على الجزء الآخر من البيانات، وهنا لا بد أن يكون حجم الجزء من البيانات الذي سيتم استخدامه لبناء النموذج يحقق الحد الأدنى من حجم العينة اللازم للانحدار الخطي.

٥- توثيق النتائج النهائية للتحليل وصياغتها لتضمينها في تقرير البحث بأسلوب علمي.

كيفية تنفيذه باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سيتم توضيح كيفية تنفيذ تحليل الانحدار الخطي المتعدد للعلاقة بين متغير تابع كمي وعدة متغيرات كمية ونوعية وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (٩-١):

يرغب باحث في مجال الرعاية الصحية في التنبؤ بقيم متغير «القدرة الهوائية القصوى» VO2max للأفراد (مل | دقيقة | كجم) - والذي يعرف على أنه كمية الأوكسجين الذي تستخلصه أنسجة الجسم من هواء الشهيق عند الجهد البدني الأقصى، وهو يعتبر مؤشراً للصحة واللياقة البدنية - من خلال مجموعة من المتغيرات المستقلة تشمل العمر (بالسنوات)، الوزن (كجم)، معدل ضربات القلب (نبضة | دقيقة)، والنوع (ذكر- أنثى). لذا قام باختيار ١٠٠ شخص عشوائياً وأجرى عليهم اختبار «القدرة الهوائية القصوى» وكانت النتائج كالتالي:

جدول رقم (٩-٢)

نتائج اختبار القدرة الهوائية القصوى لـ ١٠٠ مشارك

رقم المشارك	العمر	الوزن	النبضات	النوع	VO2	رقم المشارك	العمر	الوزن	النبضات	النوع	VO2
١	٢٧	٧٠,٤٧	١٥٠	١	٥٥,٧٩	٥١	٣٣	٦٢,٠١	١٧٢	٠	٣٦,٤٩
٢	٦٣	٥٠,٣٤	١٤٤	٠	٣٥	٥٢	٢٣	٦٨,٠٣	١٢٧	١	٦١,٧١
٣	٣٦	٨٧,٦٥	١٦٢	١	٤٢,٩٣	٥٣	٢٩	٧٩,٣٩	١٣٣	٠	٣٣,٨٤
٤	٢٦	٨٩,٨	١٢٩	٠	٢٨,٣	٥٤	٢٥	٨١,٩١	١٣٨	٠	٥١,٢
٥	٢٤	١٠٣,٠٢	١٤٣	١	٤٠,٥٦	٥٥	٣٢	٦٤	١٧٨	٠	٣٥
٦	٢٩	٧٧,٣٧	١٥٢	٠	٣٣	٥٦	٢٩	٥٩,٧٨	١٢٢	٠	٤٧,٧١
٧	٢٤	٨٢,٤٨	١٧٥	١	٤٣,٤٨	٥٧	٢١	٧١,٠٣	١٢٢	١	٥٢,٣
٨	٢٧	٧٥,٩٤	١٦٠	٠	٣٠,٣٨	٥٨	٣٤	٥٠	١٨٣	٠	٤٢
٩	٢٥	٩٧,١١	١٤٨	١	٤٠,١٧	٥٩	٢٥	٨٠,٩٨	١١٩	١	٥٥,٦٦
١٠	٢٢	٧٨,٤٢	١٢٥	٠	٣٦,٠١	٦٠	٣١	١١٢,٥٩	١١٦	١	٣٧,٣٤
١١	٣٠	٨٨,٠٢	١٥٥	١	٤٤,٢٢	٦١	٦٤	٧٣,٣٨	١١٧	٠	٤٠,٥٨
١٢	٤٥	٧٤,٤٧	١٢٣	٠	٣٨,٧٦	٦٢	٢٦	٦٢,٨٥	١٤٩	١	٦٠,٥٥
١٣	٢٥	٧٥,٩٨	١٤٧	٠	٣٣,٠٩	٦٣	٤٥	٦٨,٢٩	١٤٥	٠	٣٧,٩٣
١٤	٣٦	٥٨,٩٧	١٣٩	٠	٤٤,٨١	٦٤	٢١	٩٤,٦٨	١٣١	١	٤٤,٩٤
١٥	٢٣	١١١,٨	١٤٥	١	٣١,٩٤	٦٥	٢٨	١٠٣,٢٣	١٨٦	١	٣٥,٠١
١٦	٢٩	٧٩,٨١	١٢٨	٠	٣٤,٤٨	٦٦	٤٥	٨١,٠١	١٦٩	١	٤٥,٥٧
١٧	٣٧	٥٦,١٨	١٦٣	١	٤٧,٢٣	٦٧	٢٢	٦٩,٠٢	١٥٠	٠	٣٦,٦٣
١٨	٣٠	٨٦,١٣	١٥٦	١	٤٥,٠٦	٦٨	٢٠	٥٥	١٤١	٠	٦٠
١٩	٣٦	٨٧,٣	١٢٧	١	٥٥,١٢	٦٩	٢٢	٧٣,٤٤	١١٧	٠	٤٠,٥٢

تابع - جدول رقم (٩-٢).

رقم المشارك	العمر	الوزن	النبضات	النوع	VO2	رقم المشارك	العمر	الوزن	النبضات	النوع	VO2
٢٠	٣٦	٨٨,٥٢	١٤٧	١	٤٥,٥٨	٧٠	٣٣	٦٠,٣١	١١٩	٠	٤٧,٩٢
٢١	٢٦	٦٨,٥٢	١٤٧	٠	٣٧,٥٢	٧١	٣٨	٨٥,٩٤	١٥٣	١	٤٥,٨٤
٢٢	٣١	٧٦,٧	١٢٩	٠	٣٦,٢٧	٧٢	٣٠	٦٤,٦٥	١٤١	١	٦٠,٩٢
٢٣	٣٠	٦٢,٥١	١٤٠	١	٦٢,٥	٧٣	٥٢	٨٨,٨٣	١٠٩	١	٣٢
٢٤	٢٤	٦٠,٩	١٧٢	٠	٣٧,٠٩	٧٤	٢٢	٨٢,٤١	١٤٦	١	٤٩,٤
٢٥	٢٦	٥٣,٤٣	١٥٨	١	٤٤,٢٧	٧٥	٢٣	٦٣	١٤٩	٠	٥٠
٢٦	٢٣	٦٨,٥٢	١٣٣	٠	٤٠,٣٤	٧٦	٥٣	٧٨,٦٧	١٤٤	٠	٣١,٩٩
٢٧	٤٥	٧٢,٣٧	١٤٧	١	٥٥,١٩	٧٧	٣٣	٩٥,١	١٥٥	١	٣٩,٩
٢٨	٣٨	٨٢,١٦	١٤٤	٠	٤٩,٨٧	٧٨	٤١	٨٧,٠٥	١٤٣	١	٤٧,١٦
٢٩	٣٠	٩٠,٩٢	١٧٧	١	٣٨,٠٦	٧٩	٢٢	٦٩,٩٧	١٢١	١	٦١,٧٦
٣٠	٢١	٩٥,٥١	١٣٩	١	٤٨,١٣	٨٠	٢٦	٩٠,٦٣	١٠٠	٠	٣٣,٥٨
٣١	٣٩	٧٣,١٤	١٠٨	٠	٤٢,٤٢	٨١	٣٢	٩٢,٤٣	١٢٦	١	٤٧,٢٦
٣٢	٢٦	٦٢,٥٩	١١١	١	٤٨,٢٣	٨٢	٢٧	٧٩,٣١	١٢٩	١	٥٤,٦٧
٣٣	٢٨	٩١,٣٨	١٥١	١	٤٢,٩٦	٨٣	٢٣	٨٠,٦١	١٥٠	١	٤٩,٦٨
٣٤	٤١	٩٤,٢٩	١٤٥	١	٤٢,٥٣	٨٤	٣١	٧٤,٨٧	١٦٤	١	٥٠,١٩
٣٥	٤٠	٥٨,٠٧	١٦١	١	٤٠,٩	٨٥	٢٤	١٠١,٦٢	١٥٤	١	٣٦,١٩
٣٦	٢٥	٩٠,٢٥	١٣٢	٠	٢٧,٣٥	٨٦	٢٥	٧٤,٧	١٢٨	١	٥٧,٦٤
٣٧	٢١	٧٨,٤٥	١١١	٠	٣٨,٧	٨٧	٥٣	٨٧,٢٨	١٢٩	١	٤٩,٨٧
٣٨	٢٢	٧٤,٧٨	١٤٧	١	٦٢,١٣	٨٨	٣٣	٧٢,٥٧	١٤٨	٠	٣٤,٨٥
٣٩	٢٣	٨٥,١٣	١٥٦	١	٤٥,٦٩	٨٩	٣٢	٧٥,٨٢	١٤٤	٠	٣٣,٦٧

تابع - جدول رقم (٩-٢).

رقم المشارك	العمر	الوزن	النبضات	النوع	VO2	رقم المشارك	العمر	الوزن	النبضات	النوع	VO2
٤٠	٢٨	١٠١,٢٥	١٣٣	١	٤٠,٧٣	٩٠	٣٣	٥٣	١٥٣	٠	٤٢
٤١	٢١	٥٨,٩٤	١٥٤	١	٤١,٨٢	٩١	٢٤	٧٥,١٥	١٢٣	١	٥٨,٣٢
٤٢	٣٧	١٠١,٠٦	١٣١	١	٣٥,٦	٩٢	٤٠	١٠٣,٥٣	١٥٥	١	٤٥,٢٣
٤٣	٢١	٩٧,٧٥	١٣٦	١	٤٢,٢	٩٣	٢٣	٧٩,٦٥	١٤٩	٠	٣٠,٣٧
٤٤	٢٢	٧٢,٥	١٢٣	١	٤٥,٣	٩٤	٢٦	٧٠,٨٤	١٤٨	١	٥٥,٩١
٤٥	٢٧	٨٨,٤٥	١٧٥	١	٤٠,٠٢	٩٥	٣٨	٧٩,١٧	١٥٥	١	٤٩,٤٢
٤٦	٢٤	٧٢,٢٩	١٨٧	١	٤٧,١٧	٩٦	٣٢	١١١,٩٨	١٣٥	١	٣٣,٧٣
٤٧	٣١	٩٤,٥٩	١٣٤	١	٤٤,٤٣	٩٧	٣٣	٨٨,٠٧	١٣٠	١	٤٩,٢٢
٤٨	٢٨	١١٥,٤٢	١٣٤	١	٣٨,١٢	٩٨	٣٥	٩٤,٠٧	١٤٨	١	٤١,٩٧
٤٩	٣٠	٥١,٩٦	١٣٤	٠	٥٠,٠٥	٩٩	٤٨	٩٧,٤٦	١١٦	١	٣٥,٢
٥٠	٣١	٨٩,٣٥	١٢٩	١	٤٨,٦٤	١٠٠	٥١	٩٣,٧٦	١٢٤	١	٣٢

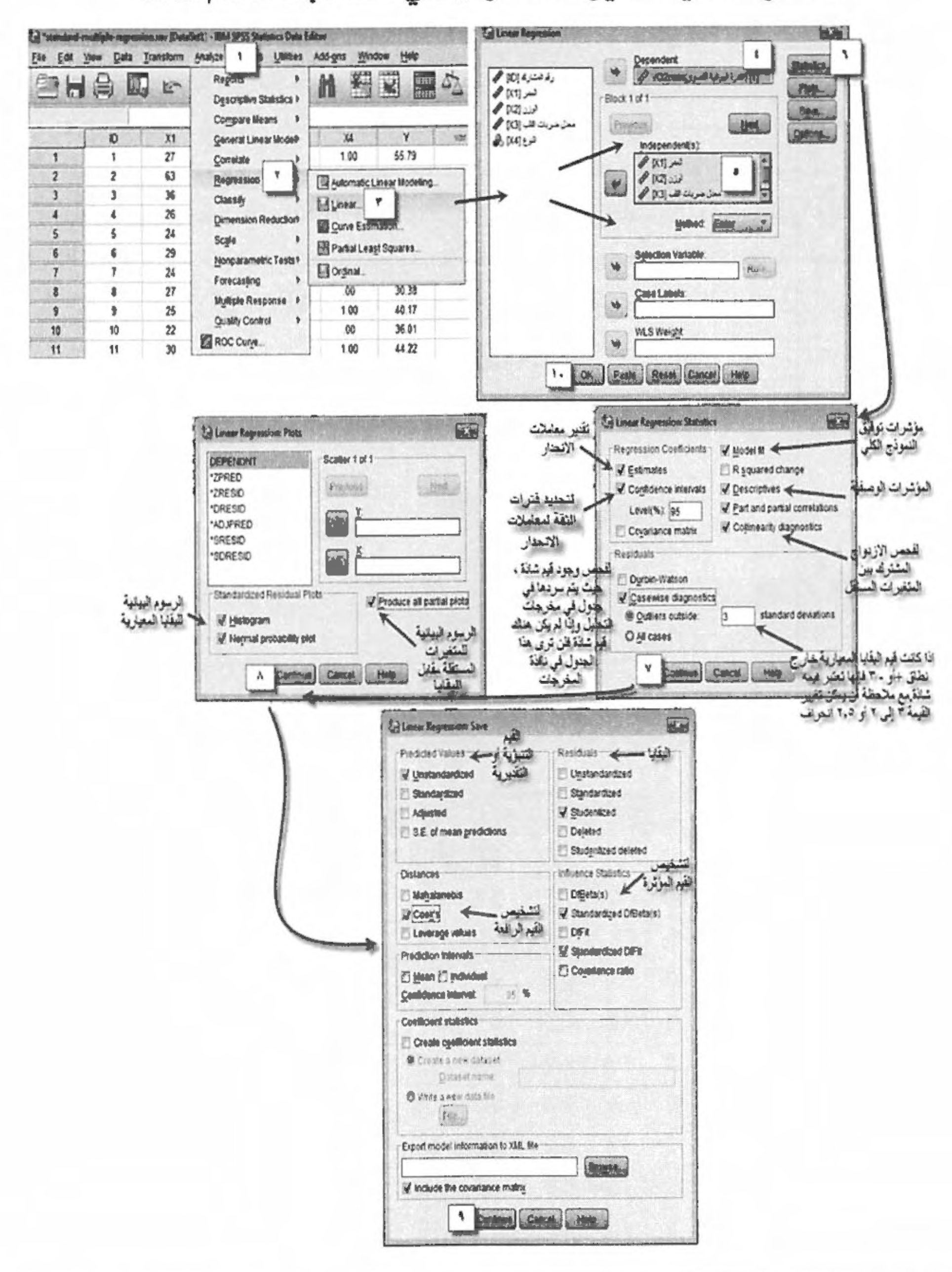
(المصدر: hpttps://statistics.laerd.com/features-tests.ph).

والمطلوب: إيجاد معادلة الانحدار الخطي المتعدد لوصف وتحليل العلاقة بين متغير القدرة القهوائية القصوى VO2max (المتغير التابع) والمتغيرات المستقلة، والتأكد من مدى صلاحيتها للتنبؤ.

خطوات الحل:

١- لإجراء تحليل الانحدار الخطي المتعدد، قم أولاً بإدخال المتغيرات والبيانات إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ١٠ كما هو موضح بالأشكال التالية:

شكل رقم (٩-٤)
خطوات تنفيذ تحليل الانحدار الخطي المتعدد باستخدام SPSS



٢- وبعــد التنفيذ يتم الحصول على متغيرات جديدة متعلقــة بالقيم التنبؤية للمتغير التابــع، والبقايا المعيارية والفروق المعيارية في القيم التنبؤية، ومعاملات الانحدار في نافذة تحرير البيانات Data Viewer - انظر الشــكل (٩-٥). وهذه المتغيرات يســتفاد منها في تشخيص مشكلة القيم الشاذة والمؤثرة بالإضافة إلى استخدامها في رســم الشــكل الانتشــاري لمتغيري البقايا المعيارية والقيــم التنبؤية. كما يتم الحصول على كم كبير من مخرجات التحليل تتضمن جداول ورسومات بيانية في نافذة المخرجات Output. وسنستعرض تلك المخرجات والتعليق عليها واستخلاص النتائج منها، ثم بعد ذلك سيتم وضع نتائج التحليل معاً وفقاً لمنهجية «جمعية علماء النفس الأمريكية APA» لتضمين نتائج التحليل في تقرير البحث أو الدراسة.

شكل رقم (٩-٥)

نافذة محرر البيانات Data Viewer بعد تنفيذ تحليل الانحدار الخطي المتعدد باستخدام SPSS

القيم التنبؤية لمتغير Y

البقايا المعيارية

الفروق المعيارية للقيم التنبؤية Standardized DfFit

الفروق المعيارية في معاملات الانحدار Standardized DfBeta

*standard-multiple-regression.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

Visible: 14 of 14 Variables

	ID	X1	X2	X3	X4	Y	PRE_1	SRE_1	SDF_1	SDB0_1	SDB1_1	SDB2_1	SDB3_1	SDB4_1
1	28	38	82.16	144	.00	49.87	32.97	3.03	.66	-.24-	.23	.29	.15	-.51-
2	54	25	81.91	138	.00	51.20	35.92	2.74	.57	.01	-.21-	.24	.01	-.45-
3	68	20	55.00	141	.00	60.00	46.74	2.40	.61	.34	-.32-	-.30-	-.06-	-.16-
4	38	22	74.78	147	1.00	62.13	51.31	1.93	.35	.13	-.19-	-.14-	.00	.19
5	92	40	103.53	155	1.00	45.23	36.33	1.61	.39	-.29-	.17	.26	.17	-.01-
6	19	36	87.30	127	1.00	55.12	46.54	1.53	.26	.07	.08	.00	-.13-	.12
7	75	23	63.00	149	.00	50.00	42.23	1.39	.29	.07	-.13-	-.07-	.05	-.14-
8	23	30	62.51	140	1.00	62.50	55.53	1.25	.28	.16	-.01-	-.22-	-.07-	.19
9	27	45	72.37	147	1.00	55.19	48.43	1.22	.28	-.01-	.20	-.12-	.01	.14
10	87	53	87.28	129	1.00	49.87	43.50	1.17	.34	-.04-	.29	.00	-.08-	.10

- **أولاً:** نقوم بفحص مصفوفة الارتباط بين المتغيرات المستقلة كما هو موضح بالشكل (٩-٦)، حيث يتضح لنا أن جميع قيم معاملات ارتباط بيرسون بين المتغيرات المستقلة أقل بكثير من ٠,٧٠ مما يعنى أنه لا يوجد ارتباط عالٍ بينها، وهذا مؤشر على عدم وجود مشكلة الازدواج الخطي بين المتغيرات المستقلة.

شكل رقم (٩-٦)

بعض المؤشرات الوصفية لمتغيرات الدراسة كما هو ظاهر في نافذة المخرجات

متوسطات المتغيرات

الإنحرافات المعيارية للمتغيرات

عدد الحالات أو حجم العينة

Descriptive Statistics

	Mean	Std. Deviation	N
VO2max(القدرة الهوائية القصوى)	43.6298	8.57131	100
العمر	31.10	9.143	100
الوزن	79.6603	15.08984	100
معدل ضربات القلب	141.99	18.285	100
النوع	.6300	.48524	100

لاحظ عدم وجود ارتباط عالي أو قوي بين المتغيرات المستقلة جميع معاملات الارتباط أقل بكثير من ٠,٧٠

Correlations معاملات ارتباط بيرسون

		VO2max(القدرة الهوائية القصوى)	العمر	الوزن	معدل ضربات القلب	النوع
Pearson Correlation	VO2max(القدرة الهوائية القصوى)	1.000	-.191-	-.307-	-.072-	.424
	العمر	-.191-	1.000	-.004-	-.061-	-.044-
	الوزن	-.307-	-.004-	1.000	-.131-	.450
	معدل ضربات القلب	-.072-	-.061-	-.131-	1.000	.108
	النوع	.424	-.044-	.450	.108	1.000
Sig. (1-tailed)	VO2max(القدرة الهوائية القصوى)	.	.028	.001	.238	.000
	العمر	.028	.	.486	.272	.332
	الوزن	.001	.486	.	.098	.000
	معدل ضربات القلب	.238	.272	.098	.	.143
	النوع	.000	.332	.000	.143	.
N	VO2max(القدرة الهوائية القصوى)	100	100	100	100	100
	العمر	100	100	100	100	100
	الوزن	100	100	100	100	100
	معدل ضربات القلب	100	100	100	100	100
	النوع	100	100	100	100	100

- **ثانياً:** نقوم بتوفيق معادلة الانحدار الخطي المتعدد، وذلك باستخدام قيم معاملات الانحدار الخطي الجزئية الموضحة بالشكل (٩-٧) من مخرجات **SPSS**:

$$\hat{y} = 87.83 - 0.156x_1 - 0.385x_2 - 0.118x_3 + 13.208x_4$$

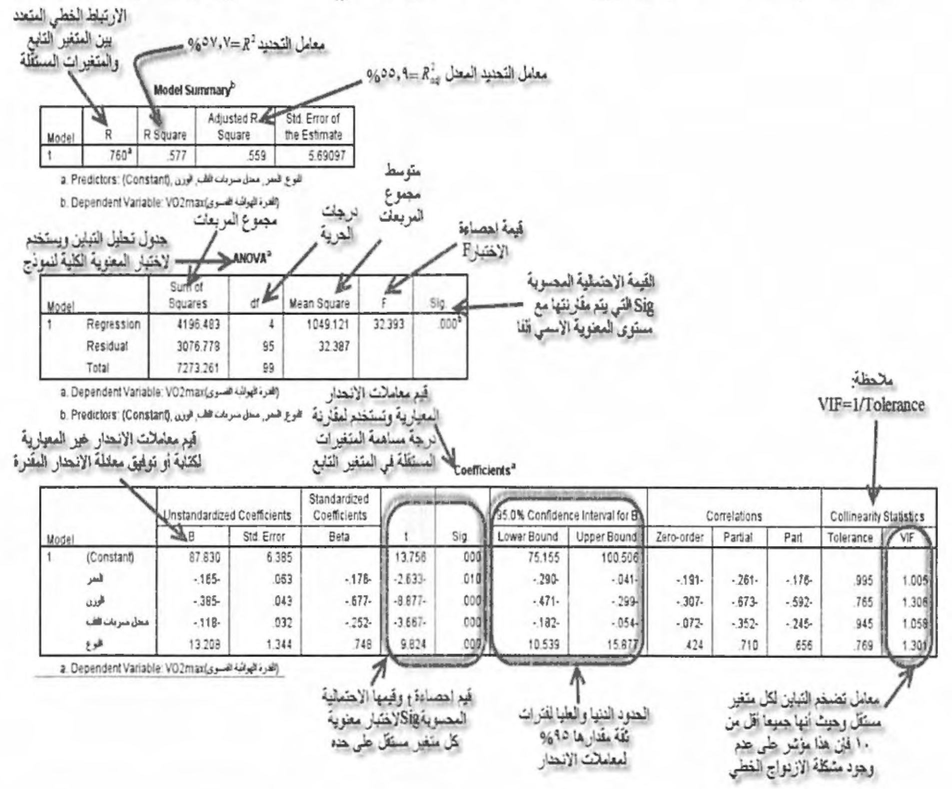
أي أن

متوسط القدرة الهوائية القصوى المقدر = ٨٧,٨٣ – ٠,١٦٥ × العمر – ٠,٣٨٥ × الوزن – ٠,١١٨ × معدل ضربات القلب + ١٣,٢٠٨ × النوع

ويمكن تفسير قيم معاملات الانحدار الجزئية في المعادلة أعلاه، فمثلاً معامل الانحدار الجزئي لمتغير الوزن يفسر على أنه لكل زيادة في الوزن بمقدار واحد كجم مع تثبيت بقية المتغيرات المستقلة في المعادلة عند قيمة معينة فإن متوسط القدرة الهوائية القصوى ينقص بمعدل ٠,٣٨٥ مل | د | كلجم. ويمكن تفسير بقية معاملات الانحدار الجزئية بنفس الطريقة.

شكل رقم (٩-٧)

إحصائيات متعلقة بتوفيق معادلة الانحدار الخطي المتعدد، وجودة توفيقها



Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.760[a]	.577	.559	5.69097

a. Predictors: (Constant), النوع, العمر, معدل ضربات القلب, الوزن

b. Dependent Variable: VO2max(القدرة الهوائية القصوى)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4196.483	4	1049.121	32.393	.000[b]
	Residual	3076.778	95	32.387		
	Total	7273.261	99			

a. Dependent Variable: VO2max(القدرة الهوائية القصوى)

b. Predictors: (Constant), النوع, العمر, معدل ضربات القلب, الوزن

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	87.830	6.385		13.756	.000	75.155	100.506					
	العمر	-.165-	.063	-.178-	-2.633-	.010	-.290-	-.041-	-.191-	-.261-	-.176-	.995	1.005
	الوزن	-.385-	.043	-.677-	-8.877-	.000	-.471-	-.299-	-.307-	-.673-	-.592-	.765	1.306
	معدل ضربات القلب	-.118-	.032	-.252-	-3.667-	.000	-.182-	-.054-	-.072-	-.352-	-.245-	.945	1.058
	النوع	13.208	1.344	.748	9.824	.000	10.539	15.877	.424	.710	.656	.769	1.301

a. Dependent Variable: VO2max(القدرة الهوائية القصوى)

- ثالثاً: القدرة التنبؤية للنموذج:

ويتــم ذلــك بفحص معامل الارتبــاط المتعدد R، ومعامل التحديــد R^2 ويتضح من الشكل (٩-٧) السابق أن:

$R = 0.76$ وتمثــل معامل الارتباط المتعدد بين القيم الفعلية والقيم التنبؤية للمتغير التابع «القدرة الهوائية القصوى» وهذه العلاقة قوية.

$R^2 = 57.7\%$ ويســمى معامل التحديد، وهو مربع معامــل الارتباط المتعدد ويعبر عنه في الصيغــة المئوية وتتراوح قيمته من ٠٪ (انعدام القــدرة التنبؤية للنموذج) إلى ١٠٠٪ (قدرة تنبؤية تامة). ويفســر إحصائياً على أنه نســبة التباين في المتغير التابع الذي يفسر بواســطة التباين في المتغيرات المستقلة مجتمعة. ويستخدم كمؤشر على جودة توفيق معادلة الانحدار المقدرة لبيانات العينة. ويعاب على R^2 أنها متحيزة دائماً للأعلــى فهي تزيد بزيادة عدد المتغيرات المســتقلة في النمــوذج حتى وإن كان المتغير المســتقل المضــاف للنموذج غير مفيد في تفســير المتغير التابــع ولا يمثل أي إضافة حقيقية لدقة النموذج. وفي هذا التطبيق العملي فإن ٥٧,٧٪ من التباين في «القدرة الهوائية القصوى» يفســره التباين في المتغيرات المستقلة «العمر، الوزن، معدل ضربات القلب، والنوع» مجتمعة.

$R^2_{adj} = 55.9\%$ يســمى معامــل التحديد المعدل ويفســر بنفس طريقة R^2 ويفضل على R^2 في الدراسات والبحوث حيث إنه غير متحيز للأعلى بمعنى أنه لا تزيد قيمته مع أي زيادة في عدد المتغيرات المســتقلة بل إن قيمتــه تنقص إذا تم إضافة متغير أو متغيرات مســتقلة غير مفيدة في تفسير المتغير التابع ويستخدم مؤشراً لتقدير حجم التأثيــر Effect Size. وفي هــذا التطبيق فإن القدرة التفســيرية أو التنبؤية لمعادلة الانحدار الخطي المتعدد المقدرة تعتبر كبيرة، وذلك حسب تصنيف Cohen (1992).

- رابعاً: اختبار المعنوية الإحصائية الكلية لمعادلة الانحدار الخطي المقدرة:

ولاختبار معنوية تلك المعادلة يتم صياغة الفرضيتين التاليتين:

- فرضيــة العدم H_0: جميع معاملات الانحدار الخطي الجزئية = الصفر (أي جميع المتغيرات المستقلة مجتمعة لا ترتبط بالمتغير التابع ارتباطاً معنوياً).
- الفرضية البديلة H_1: يوجد معامل انحدار خطي جزئي واحد على الأقل ≠ الصفر (أي أحد المتغيرات المستقلة يرتبط بالمتغير التابع ارتباطاً معنوياً).

وحيــث إن قيمــة $Sig = 0.000$ (تكتب في تقرير البحــث $p < 0.001$) الاحتمالية المحسوبة لإحصائية الاختبار $F(4,95) = 32.39$ كما يوضحه جدول تحليل التباين في الشكل (٩-٧) أقل بكثير من $\alpha = 0.05$ فإن الباحث يرفض فرضية العدم ويتوصل إلى أن بيانات العينة تؤيد وجود علاقة ارتباط معنوية بين المتغير التابع وبين أحد المتغيرات المستقلة على الأقل.

- خامساً: اختبار المعنوية الجزئية لكل متغير مستقل على حدة:

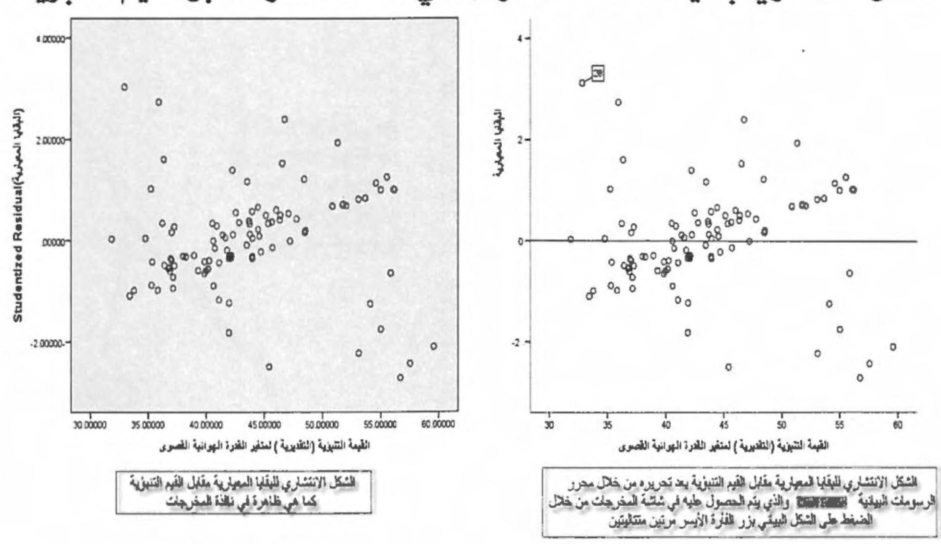
مــن الملاحظ من جدول «المعامــلات Coefficients» في الشــكل (٩-٧) أن جميع المتغيرات المســتقلة (معاملات الانحدار الجزئية) مرتبطة معنوياً بالمتغير التابع، وذلك لكون القيمة الاحتمالية المحسوبة Sig. المقابلة لإحصاءة الاختبار t لكل منها أقل من $\alpha = 0.05$ لذا سيتم الإبقاء على جميع تلك المتغيرات في معادلة الانحدار المقدرة.

- سادساً: تشخيص النموذج:

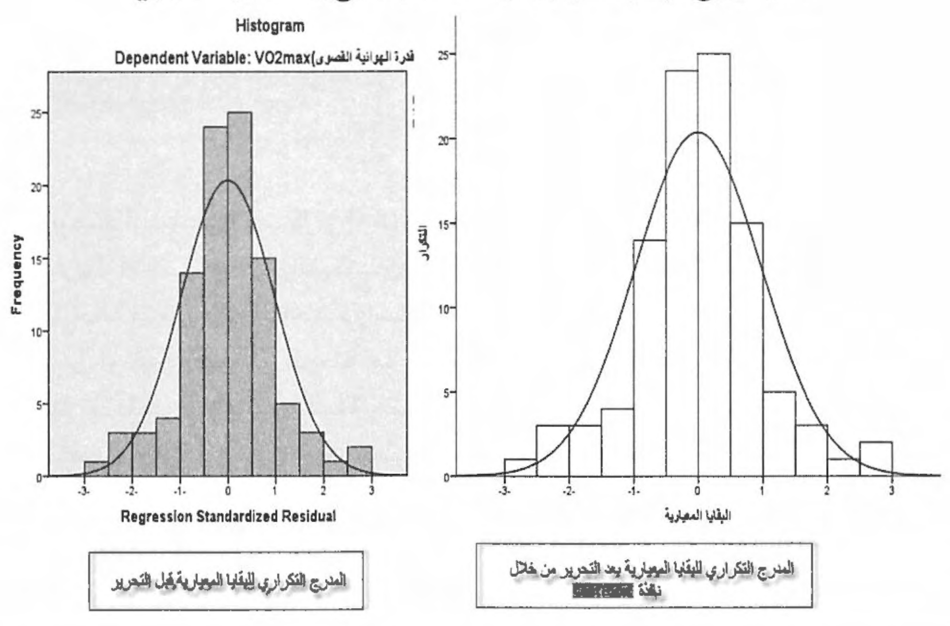
ســنعمل الآن على فحص مدى تحقق شــروط الانحدار الخطي المتعدد من (١-٦) السابق ذكرها كما يلي.

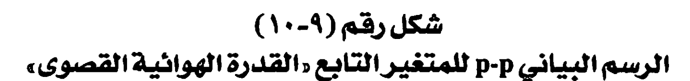
وكما ذكرنا سابقاً فإن الشكل الانتشاري للبقايا المعيارية على المحور الرأسي مقابل القيم التنبؤية للمتغير التابع على المحور الأفقي يتم اســتخدامه في تشخيص الشروط الســابق ذكرهــا. فإذا كان نمط انتشــار البيانات حول الخــط الأفقي - الذي يمر بمتوســط البقايا المعيارية المســاوي للصفر- عشــوائياً كما في الشكل ٩-٢ (أ)، فإنه يمكن القول بأن الشــروط الخطية، والتوزيع الطبيعي، وتجانس التباين واســتقلالية الخطأ العشوائي متحققة. وفي هذا التطبيق فإنه يمكن استنتاج بأن الشروط الأربعة الآنفة الذكر متحققة - انظر الشكل (٩-٨). كما أن الشكلين (٩-٩) و(٩-١٠) يدلان على أن توزيع حد الخطأ العشوائي طبيعي.

شكل رقم (٩-٨)
الشكل الانتشاري لبقايا معادلة الانحدار الخطي المتعدد المقدرة مقابل القيم التنبؤية



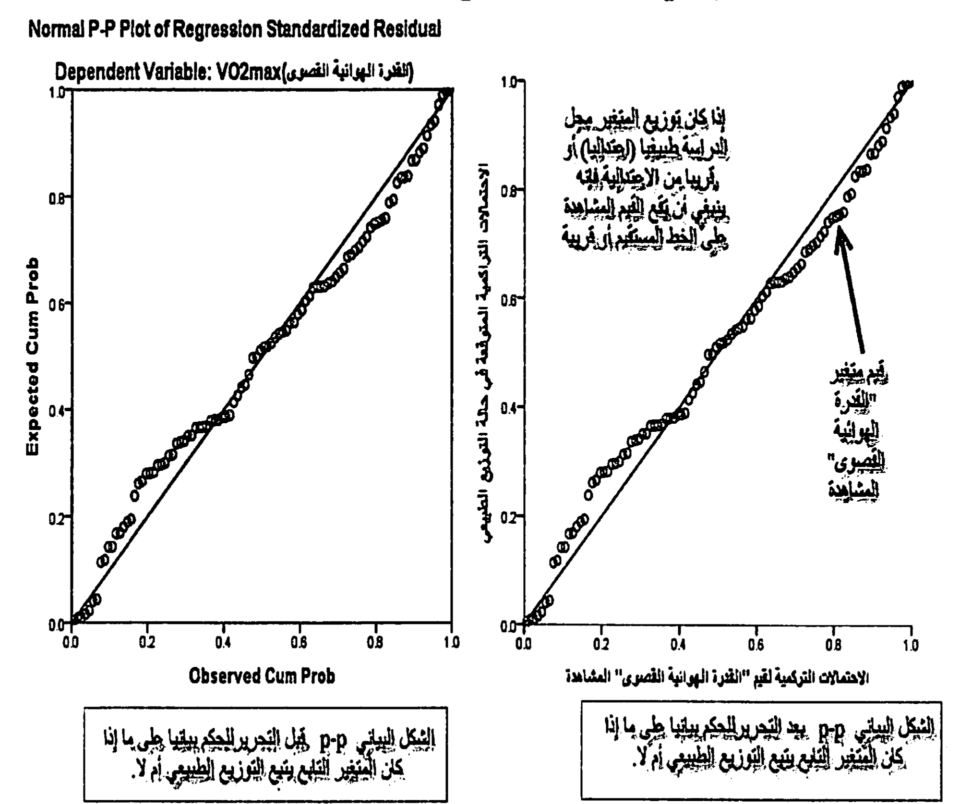
شكل رقم (٩-٩)
شكل توزيع البقايا المعيارية باستخدام المدرج والمنحنى التكراري





شكل رقم (٩-١٠)

الرسم البياني p-p للمتغير التابع «القدرة الهوائية القصوى»



ويتضح أيضاً من الشــكل (٩-٨) وجود قيمة شــاذة واحدة فقــط خارج النطاق أو الشــريط الأفقي المحدد بالقيمتــين 3 ± وهي تمثل الحالة أو المشــارك رقم ٢٨، وفي هذه الحالة ينبغي على الباحث دراســة تلك القيمة الشــاذة. فقد تكون نتيجة تسجيل خاطــئ أو غير دقيق في مرحلة جمع البيانات، أو قد تكون إدخالاً خاطئاً للبيانات في مرحلة تفريغها، أو قد تكون قيمة حقيقية. فإذا كانت نتيجة تســجيل أو إدخال خاطئ فيتم تعديلها، أما إذا كانت قيمة فعلية فإنه يتم دراستها، وذلك من حيث درجة تأثيرها في النمــوذج. فإذا كانت غير مؤثرة في نموذج الانحدار المقدر فإنه يتم الإبقاء عليها. أمــا إذا كانت مؤثــرة في دقة التنبؤ لمعادلة الانحدار المقدرة فقد يلجأ الباحث إلى عدة خيارات منها: إجراء تحويل خطي على المتغير المسبب للقيمة الشاذة أو يستخدم أسلوب

الانحـــدار المانع Robust مثلاً. أو قد يلجأ الباحث إلى حذف الحالة إذا لم يكن ذلك ســـيؤدي إلى تحيز أو إخلال بنتائج الدراســـة، أو قد ينشئ الباحث معادلتي انحدار مع وبدون القيمة الشـــاذة. وفي هذا التطبيق سنرى ما إذا كانت تلك القيمة الشاذة مؤثرة أم لا، وذلـــك من خلال إجراء مزيد من الفحص لمؤشـــرات القيـــم المؤثرة. وكما يتضح من الشـــكل (٩-١١) أدناه، فإن أعلى قيمة لمؤشر Cook's D أقل من الواحد الصحيح، وهذا يعني عدم وجود قيم مؤثرة. كما أن القيم العظمى والصغرى كما يوضحه الشكل (٩-١٢) لمؤشـــرات الفروق المعيارية فـــي القيم التنبؤيـــة Standardized DfFits وقيم الانحدار الخطي Standardized DfBetas أقل من الواحد الصحيح أيضاً.

شكل رقم (٩-١١)
إحصائيات البقايا لمعادلة الانحدار المقدرة

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	31.8360	59.5796	43.6298	6.51066	100
Std. Predicted Value	-1.811-	2.450	.000	1.000	100
Standard Error of Predicted Value	.818	2.359	1.237	.298	100
Adjusted Predicted Value	31.8212	60.7639	43.6755	6.57257	100
Residual	-14.91621-	16.89905	.00000	5.57482	100
Std. Residual	-2.621-	2.969	.000	.980	100
Stud. Residual	-2.709-	3.032	-.004-	1.010	100
Deleted Residual	-15.93976-	17.62248	-.04567-	5.93294	100
Stud. Deleted Residual	-2.806-	3.174	-.004-	1.029	100
Mahal. Distance	1.057	16.016	3.960	2.602	100
Cook's Distance	.000	.141	.013	.026	100
Centered Leverage Value	.011	.162	.040	.026	100

a. Dependent Variable: VO2max(القدرة الهوائية القصوى)

القيمة العظمى للبقايا المعيارية تتجاوز ٣

حيث أن أعلى قيمة لمؤشر Cook's D هو ١٤١,٠ وهو أقل من الواحد الصحيح، فإننا نستنتج أنه لا يوجد قيم مؤثرة سلبا في معادلة الانحدار المقدرة

شكل رقم (٩-١٢)
القيم العظمى والصغرى للفروق المعيارية في القيم التنبؤية وقيم معاملات الانحدار

القيم الصغرى ← Minimum
القيم العظمى ← Maximum
الفروق المعيارية للقيم التنبؤية ← Standardized DFFIT
الفروق المعيارية في قيم معاملات الانحدار ← Standardized DFBETA

Descriptive Statistics

	N	Minimum	Maximum
Standardized DFFIT	100	-.86362-	.65668
Standardized DFBETA Intercept	100	-.61254-	.34471
Standardized DFBETA X1	100	-.60278-	.39323
Standardized DFBETA X2	100	-.29934-	.57549
Standardized DFBETA X3	100	-.16080-	.49474
Standardized DFBETA X4	100	-.50921-	.20177
Valid N (listwise)	100		

أمـا فيما يتعلق بمشـكلة «الازدواج الخطي» فهناك عدة طرق لتشـخيصها منها: فحص مصفوفة الارتباط الخطي للمتغيرات المسـتقلة، ومما سـبق تم اسـتنتاج أنه لا يوجد ارتباط عالٍ بين المتغيرات المستقلة. ومن المؤشرات الإحصائية التي تستخدم لغرض الكشـف عن الازدواج الخطي معامل تضخم التباين VIF للمتغيرات المستقلة. وكما يتضح من الشـكل (٩-٧)، فإن جميع معاملات تضخـم التباين VIF للمتغيرات المستقلة أقل من ١٠، وهذا مؤشر على عدم وجود مشكلة الازدواج الخطي.

- سابعاً: ملخص النتائج وتوثيقها:

وبعد بناء النموذج وتحليل قدرته التنبؤية وتشخيص شروطه، سنقوم الآن بتلخيص كل ما سـبق مناقشته في هذا التطبيق استعداداً لتوثيق نتائج تحليل الانحدار الخطي المتعدد وفقاً لأسلوب APA كالتالي:

لقد تم استخدام الانحدار الخطي المتعدد للتنبؤ بمتغير «القدرة الهوائية القصوى» من متغيرات العمـر، الوزن، معدل نبضات القلب، والنوع. وقد تم التحقق من تحقيق شـروط الخطية، التوزيع الطبيعي للأخطاء العشـوائية، استقلاليتها، تجانس التباين لهـا، وعدم وجود قيم شـاذة مؤثرة. وقد اتضح أن هذه المتغيرات المسـتقلة مجتمعة

ترتبط معنوياً بمتغير «القدرة الهوائية القصوى» وتفسر ٥٥,٩٪ من التباين فيه، $F (4,95) = 32.393, P < 0.001, Adj. R2 = .559$. كما أن كل المتغيرات المستقلة ساهمت معنوياً في التنبؤ، $p < 0.05$. والجدول (٩-٣) يحتوي على معاملات الانحدار وأخطائها المعيارية.

جدول رقم (٩-٣)
ملخص تحليل الانحدار المتعدد

المتغير	B	SEB	β
الثابت	٨٧,٨٣٠	٦,٣٨٥	
العمر	- ٠,١٦٥*	٠,٠٦٣	- ٠,١٧٦
الوزن	- ٠,٣٨٥*	٠,٠٤٣	- ٠,٦٧٧
معدل نبضات القلب	- ٠,١١٨*	٠,٠٣٢	- ٠,٢٥٢
النوع	١٣,٢٠٨*	١,٣٤٤	٠,٧٤٨

*: $P < 0.05$، B = معاملات الانحدار الخطي الجزئية، SEB = الخطأ المعياري لمعامل الانحدار، β = معاملات الانحدار المعيارية.

ملاحظة:

إذا أراد الباحث إضافة جدول تحليل التباين في تقرير بحثه، فإنه يتم عرضه كالتالي:

جدول رقم (٩-٤)
جدول تحليل التباين

مصدر التباين	درجات الحرية	مجموع المربعات	متوسط المربعات	F	P-value
الانحدار	٤	٨٤١٩٦,٤	١٠٤٩,١٢	٣٢,٣٩٣	< 0.001
البقايا	٩٥	٣٠٧٦,٧٨	٣٢,٣٨٧		
المجموع	٩٩	٧٢٧٣,٢٦			

تمرين تطبيقي (٩-١):

الجدول التالي يحتوي على بيانات ٤٠ فرداً على المتغيرات التالية (Ho, 2006).

X1: أحداث الحياة، مقاس على مقياس أحداث الحياة Life events scale.

X2: المتاعب، مقاس على مقياس المتاعب Hassles scale.

X3: الدعم الاجتماعي، مقاس على مقياس الدعم الاجتماعي Social support scale.

Y: الاكتئاب، مقاس على مقياس الاكتئاب Depression scale.

الحالة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
X1	١٢٠	١٨٥	١٥٠	١١٠	٩٤	١٥٦	٥٥	١٣٢	١٤١	١٥٧
X2	٤٦	٤٨	٦٠	٥٥	٤١	٥٤	٣١	٣٣	٣١	٥٩
X3	١٠	٣٣	١٠	٢٣	٤٧	٤٠	٤٨	٤١	٣٢	٣٨
Y	١٥٠	١٦٠	٧٠	١٥٠	٩٠	١٣٠	١٢٠	١٢٠	١٥٠	١٦٠
الحالة	١١	١٢	١٣	١٤	١٥	١٦	١٧	١٨	١٩	٢٠
X1	٢٠٢	١٠٧	١٤٨	١٨٠	٨٨	١٠٨	١١١	١٢٢	١٦٢	١٥٠
X2	٥٣	٥١	٣٩	٤٣	٤٣	٤٣	٤٠	٤٦	٥٥	٥١
X3	١٣	٣٩	٣٧	١٣	٥١	٣٠	٥	٦٤	٦٠	٢٨
Y	١٨٠	١٢٠	١٤٠	١٢٠	١٠٠	١١٠	١٢٠	٨٠	١٤٠	١٢٠
الحالة	٢١	٢٢	٢٣	٢٤	٢٥	٢٦	٢٧	٢٨	٢٩	٣٠
X1	١٤٨	١٠٣	١٥٥	١٣٤	١١٦	١٥٩	١٢٢	١٤٩	٨٦	١٠٠
X2	٤٥	٥٠	٥٧	٤٦	٣٦	٦٠	٤٦	٣٦	٤٢	٣٦
X3	٤٧	١٨	١٩	٤٥	٤٦	١٨	٣٢	٣٩	٣٩	٤٤
Y	٩٠	١٣٠	١٤٠	١٣٠	٩٠	١٢٠	٩٠	٩٠	٨٠	١١٠
الحالة	٣١	٣٢	٣٣	٣٤	٣٥	٣٦	٣٧	٣٨	٣٩	٤٠
X1	١٢٧	١٣٠	٩٥	٥٢	٩١	١٤٨	١٥٠	١٤٣	٢٩	١٧٩
X2	٥٠	٥٤	٥٦	٣٣	٤٥	٤٣	٣٩	٤٥	٣٧	٥٥
X3	١٥	٤٩	٣٣	٤٩	٤١	٢٠	٢١	٢٤	٣٩	١٠
Y	١٥٠	١٠٠	٩٠	١١٠	١١٠	١١٠	٩٠	١٦٠	١٤٠	١٥٠

والمطلوب: إيجاد معادلة الانحدار الخطي المتعدد لوصف وتحليل العلاقة بين متغير الاكتئاب (Y) والمتغيرات المستقلة (X1, X2, X3)، والتأكد من مدى صلاحيتها للتنبؤ.

الفصل العاشر
التحليل العاملي Factor Analysis

مقدمة:

إن التحليل العاملي هو أحد أساليب التحليل الإحصائي المتقدمة لتحليل العلاقة بين المتغيرات المتعددة. ويهدف بشكل رئيسي إلى تحليل وفحص هيكل أو بنية Structure الارتباطات بين العديد من المتغيرات كي نرى ما إذا كان باستطاعتنا وصف تلك العلاقات بين تلك المتغيرات من خلال استخلاص مجموعة أصغر من المتغيرات تسمى «العوامل Factors أو المكونات Components». والتحليل العاملي لا يحتاج إلى التفريق أو التمييز بين المتغيرات التابعة والمستقلة (Field, 2007). ومن استخداماته تقليص عدد المتغيرات الكبير (وليكن ٢٥ متغيراً مثلاً) لوصف ظاهرة أو خاصية معينة إلى عدد أقل من المتغيرات أو العوامل الرئيسية (ولتكن ٥ عوامل مثلاً) كافية لوصف تلك الظاهرة أو الخاصية. حيث إنه مع العدد الكبير من المتغيرات يصعب على الباحث وصف الظاهرة بشكل واضح ومفيد يستطيع من خلاله الوصول إلى استنتاجات مثمرة ومفيدة. إذن يمكن القول بأن التحليل العاملي يعمل على البحث عن العوامل Factors أو الأبعاد Dimensions الرئيسية المكونة لمجموعة من المتغيرات Variables المرتبطة.

أنواع التحليل العاملي:

هناك نوعان رئيسيان للتحليل العاملي (2004) Thompson هما:

١- التحليل العاملي الاستكشافي Exploratory Factor Analysis. وكما يتضح من تسميته فإنه يستخدم لاستكشاف العوامل المكونة لمجموعة من المتغيرات دون فرض هيكل محدد مسبقاً من قبل الباحث لتوزيع المتغيرات على العوامل حسب درجة ارتباط المتغيرات بتلك العوامل. وللتحليل العاملي الاستكشافي العديد من الطرق لاستخلاص أو بناء العوامل من المتغيرات.

٢- التحليل العاملي التوكيدي Confirmatory Factor Analysis. وكما يتضح من تسميته فإنه يستخدم لتأكيد أن توزيع المتغيرات على العوامل يتوافق مع التوزيع النظري المحدد مسبقاً من قبل الباحث للمتغيرات على العوامل بناء على نظرية

سابقة أو أدبيات البحث. وفيه يقوم الباحث باختبار العلاقة المفترضة بين المتغيرات والعوامل إحصائياً.

التحليل العاملي الاستكشافي Explaratory Factor Analysis:

نظراً لكثرة استخدام التحليل العاملي الاستكشافي في الدراسات والبحوث الاجتماعية فإننا سنقتصر في هذا الكتاب على الحديث عن التحليل العاملي الاستكشافي (أو اختصاراً التحليل العاملي وهو الاسم الشائع بين الباحثين) وكيفية تنفيذه خطوة بخطوة باستخدام SPSS.

استخداماته:

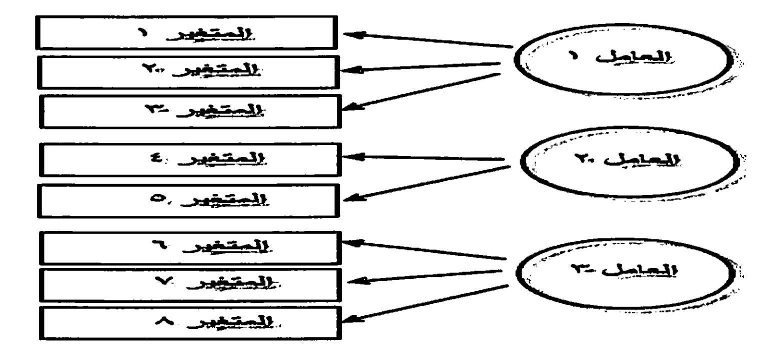
من استخدامات التحليل العاملي (Field, 2007) ما يلي:

١- فهم الهيكل البنائي لمجموعة من المتغيرات المتعلقة بظاهرة معينة مثل الذكاء، الولاء ... إلخ. والشكل (١٠-١) يصور للقارئ بعض المتغيرات حول ظاهرة معينة والعوامل التي نشأت منها أو ترتبط بها تلك المتغيرات.

٢- بناء الاستبانات لقياس مفهوم معين مثل تصميم استبانة لقياس الولاء للمنظمة.

٣- تقليص عدد المتغيرات إلى عدد أقل من المتغيرات أو العوامل الرئيسية الكافية لفهم وتفسير الظاهرة محل الدراسة بدرجة عالية من المصداقية والثقة واستخدامها في تحليلات إحصائية لاحقة إن دعت الحاجة إلى ذلك مثل تحليل الانحدار، وتحليل التباين، وغير ذلك من أساليب التحليل الإحصائي الأخرى.

شكل رقم (١٠-١)
توزيع المتغيرات على العوامل



وقد نشأ مفهوم التحليل العاملي بداية واستخدم في مجال علم النفس. ثم انتقل بعد ذلك إلى الميادين الأخرى مثل الإحصاء والجودة والاقتصاد والتسويق والإدارة والعلوم الطبيعية والطبية وغير ذلك من المجالات الأخرى، وأخذ تطبيقات علمية وعملية مختلفة. ففي مجال الإحصاء مثلاً، يمكن استخدام التحليل العاملي لمعالجة مشكلة الازدواج الخطي (Colinearity (or Multicolinearity. وقد سبق الإشارة في الفصل الأول من الكتاب إلى أن التحليل العاملي يستخدم في دراسة صدق بناء الأداة أو الاستبانة المصممة لقياس ظاهرة أو مفهوم ما.

المفاهيم والأساسيات الهامة في التحليل العاملي:

المتغيرات أو العناصر العكسية:

هي المتغيرات التي تأخذ قيما معاكسة في الاتجاه لما عليه الاتجاه العام للمتغيرات الأخرى في الاستبانة أو أداة القياس. فمثلاً في استبانة مكونة من مجموعة من العناصر أو الأسئلة أو البنود مقاسة على مقياس ليكرت الخماسي المتدرج من ١ إلى ٥، حيث ١ تعني غير موافق تماماً، و٥ تعني موافق تماماً، إذا كانت تلك الاستبانة تحوي عنصراً مصاغ بالصيغة السلبية مثل «تعامل مقدم الخدمة غير جيد» فإن الإجابة بـ ٥ والتي تعني موافق تماماً تساوي الإجابة بـ ١ والتي تعني غير موافق تماماً فيما إذا كان العنصر مصاغاً بالطريقة الإيجابية «تعامل مقدم الخدمة جيد». وفي مثالنا هذا يتم إعادة ترميز قيم العناصر العكسية من خلال طرح قيم ذلك العنصر من العدد ٦. أي أن قيم العنصر أو المتغير الجديد بعد إعادة ترميزه هي:

قيمة العنصر أو المتغير بعد إعادة الترميز = ٦- قيمة العنصر أو المتغير في وضعه الحالي وفي مثالنا السابق يكون: قيمة «تعامل مقدم الخدمة غير جيد» بعد إعادة الترميز = ٦- قيمة «تعامل مقدم الخدمة غير جيد»

وتجدر الإشارة إلى أن إعادة الترميز ضروري عندما يريد الباحث حساب معامل الثبات للاستبانة لأن البنود العكسية تؤثر على قيمة معامل الثبات، أما إذا أراد الباحث فقط استخدام التحليل العاملي، وذلك لغرض تقليص عدد البيانات مثلاً فإنه لا يلزمه إعادة الترميز، إذ لن يؤثر ذلك على درجة ارتباط البند بالعامل.

مصفوفة الارتباط Correlation Matrix:

وهـــي مصفوفة مربعة متماثلة يرمز لها بالرمز R وتحتوي على معاملات الارتباط الخطـــي (غالباً معامل ارتباط بيرســـون) بـــين كل متغيرين من المتغيـــرات. العناصر القطرية تساوي الواحد الصحيح وهو ارتباط المتغير مع نفسه، والعناصر غير القطرية تحتوي على قيم بين -١ و+١ تمثل الارتباط الخطي بين كل متغيرين مختلفين. فمثلاً مصفوفة الارتباط المربعة لثلاثة متغيرات X1, X2, X3 تأخذ الشكل التالي:

$$R = \begin{bmatrix} 1 & r_{x1x2} & r_{x1x3} \\ & 1 & r_{x2x3} \\ & & 1 \end{bmatrix}$$

ويجب أن يكون قياس المتغيرات المستخدمة رتبياً على الأقل لتنفيذ التحليل العاملي، مع إمكانية اســـتخدام متغيرات ثنائية القيم Dichotomous بعد إعادة ترميزها بقيم عددية مثل ١ و٢ (Ho, 2006). وقد ذكر Ho بأنه إذا كانت كل المتغيرات اسمية ثنائية Nominal Variables Dichotomous فإنه من الأفضل اســـتخدام نوع آخر من أنواع التحليل العاملي مثل التحليل العاملي البوليني Boolean Factor Analysis.

ويعتمد التحليل العاملي على مصفوفة التباين المشـــترك Covariance أو مصفوفة الارتباط بين المتغيرات وتصنيف تلك المتغيرات في مجموعات معينة حسب درجة ارتباط بعضها ببعض للوصول إلى العوامل الكامنة Latent Variables المكونة لتلك المتغيرات.

ملاحظة:

إذا كانت جميع العناصر غير القطرية تســـاوي الصفـــر وجميع العناصر القطرية تســـاوي الواحد الصحيـــح فإن المصفوفة في هـــذه الحالة تســـمى مصفوفة الوحدة Identity Matrix.

محددة مصفوفة الارتباط The Determinant of Correlation Matrix:

مؤشـــر رياضي لقياس كمية تباين المتغيرات في المصفوفة ذات العلاقة. ويستخدم فـــي التحليل العاملـــي معياراً لاختبار مـــا إذا كانت مشـــكلة الازدواج الخطي المتعدد Multicolinearity موجودة بين متغيرين على الأقل.

التشبع Loading:

وهو عبارة عن معامل الارتباط بين المتغير والعامل.

الجذور الكامنة Eigenvalues:

مؤشـــر إحصائي يســـتخدم في التحليل العاملي لمعرفة كمية التباين في مصفوفة الارتباط للمتغيرات التي يفســـرها العامل. وتستخدم الجذور الكامنة معياراً لتحديد عدد العوامل الكافية والممثلة التي ينبغي اســـتخلاصها من المتغيرات. فمثلاً إذا كانت قيمـــة الجـــذر الكامن لعامل معين تســـاوي ١ فـــإن هذا يعني أن تفســـير ذلك العامل لكمية التباين في البيانات تســـاوي نفس الكمية من التباين في البيانات التي تفسرها المتغيرات المرتبطة معنوياً بذلك العامل مجتمعة.

الشيوع Communality:

ويعني شـــيوع المتغير على العوامل وهو نسبة التباين في المتغير المفسر بواسطة (أو الذي يعزى إلى) العوامل المشتركة (يماثل دور R^2 في الانحدار الخطي المتعدد). وقيمة شـــيوع أي متغير على العوامل المشتركة تساوي مجموع مربعات تشبع هذا المتغير على العوامل. ويستخدم الشيوع في توزيع المتغيرات على العوامل. ويقترح Larose (2006) أن لا يقل الحد الأدنى لشـــيوع المتغيـــر عن ٠,٥٠، ويـــرى Tabachnick and Fidell (2007) أن يتم حذف المتغير إذا كان شيوعه أقل من ٠,٢٠.

درجات العوامل Factor Scores:

هي عبارة عن قياســـات أو قيم العوامل المستخلصة لكل مفردة من مفردات العينة، فعندما نصل في نهاية التحليل العاملي إلى العوامل المشـــتركة النهائية المستخلصة من المتغيرات وقبولها، فإنه يمكن اســـتخدامها في توليد ما يســـمى بدرجات العوامل لكل مفـــردة من مفردات العينة، ومن ثم اســـتخدام تلك العوامل ودرجاتها في أي تحليلات إحصائية لاحقة.

وهناك عدة طرق للحصول على درجات العوامل DiStefano, Zhu and Mindrila (2009) منها:

- طريقـــة الانحدار الخطي Regression Approach. قد ينتج عنها أن تكون درجات العوامل مرتبطة حتى وإن كانت العوامل متعامدة أو مستقلة Orthogonal.
- طريقة بارتلت Bartlett Method. قد يحدث أن تكون درجات العوامل مرتبطة حتى وإن كانت العوامل متعامدة أو مستقلة.
- طريقة أندرســـون – روبن Anderson-Rubin. وتتميز بأنها تحافظ على أن تكون درجات العوامل غير مترابطة عندما تكون العوامل متعامدة أو مستقلة.

وهذه الطرق الثلاث السابقة متوفرة ببرنامج SPSS ولا يوصى باستخدامها في الواقع العملي لأن تلك الطرق المختلفة تعطي نتائج متفاوتة بشكل كبير.

- جمع قياسات أو قيم المتغيرات لكل عامل Sum Scores by Factor. وهذه الطريقة أفضل من سابقاتها والموصى بها من قبل الباحثين والمختصين لأنها أسهل في الحساب والتفسير، وتحافظ على التباين في البيانات. وينبغي ملاحظة أن يتم تحويل المتغيرات إلى درجات معيارية Standardized Scores (أي طرح قيم المتغير من المتوسط الحسابي له ثم قسمة الناتج على الانحراف المعياري للمتغير) في حال اختلاف وحدات القياس للمتغيرات أو في حال تباينها الكبير. بالإضافة إلى ذلك، إذا كان أحد المتغيرات الذي ينتمي لعامل ما عكسياً، فإنه يجب طرح قيمه من مجموع قيم المتغيرات التي تنتمي لنفس العامل. وهذه الطريقة لا يوفرها SPSS آلياً وإنما يجب على الباحث أو المستفيد حسابها من خلال قائمة Compute في حال رغب في ذلك.

الافتراضات أو الشروط Assumptions لتنفيذ التحليل العاملي الاستكشافي:

١- أن يكون قياس المتغيرات المستخدمة رتبياً على الأقل لتنفيذ التحليل العاملي، مع إمكانية استخدام متغيرات ثنائية القيم Dichotomous بعد إعادة ترميزها بقيم عددية مثل ١ و٢ (Ho, 2006). وقد ذكر Ho بأنه إذا كانت كل المتغيرات اسمية Nominal Variables فإنه من الأفضل استخدام نوع آخر من أنواع التحليل العاملي مثل التحليل العاملي البوليني Boolean Factor Analysis.

٢- أن تكون العينة عشوائية.

٣- أن يكون حجم العينة كافياً، وذلك لتقدير الارتباط بين المتغيرات بشكل موثوق. ولقد تباينت الآراء حول حجم العينة الملائم لإجراء التحليل العاملي. وقد سبق التطرق للحديث عن الحد الأدنى لحجم العينة للتحليل العاملي في الفصل الثالث من هذا الكتاب، حيث يرى بعضهم استخدام ١٠ أو ١٥ وحدة أو مفردة لكل متغير. وقد ذكر Bartlett, Kotrlik &Higgins (2001) أنه ينبغي أن لا يقل حجم العينة في التحليل العاملي عن ١٠٠ وحدة. وفي المقابل اقترح Tabachnick and Fidell (2007) ٣٠٠ وحدة على الأقل. كما قدم Comrey and Lee (1992) توصية بخصوص حجم العينة لتنفيذ التحليل العاملي، فقد أوردا بأن ٥٠ وحدة تعتبر ضعيفة جداً، ١٠٠ وحدة إلى ٢٠٠ وحدة مقبولة، ٣٠٠ وحدة جيدة، ٥٠٠ وحدة

جيدة جــداً، ١٠٠٠ وحدة فأعلى ممتازة. كما أن Field (2005) قد تطرق لحجم العينة الملائم للتحليل العاملي واســتعرض الأدبيات حول ذلك وخلص إلى أن ٣٠٠ وحــدة على الأقل تعتبر مناســبة لتنفيــذ التحليل العاملي. وتجدر الإشــارة إلى أن هناك مؤشــراً أو معيــاراً إحصائياً كمياً خاصاً للحكم علــى ما إذا كان حجم العينة كافياً لتنفيذ التحليل العاملي أم لا (Kaiser, 1970)، وهذا المؤشــر يســمى KMO لملاءمة العينة Kaiser-Meyer-Olkin (KMO) for Sampling adequacy وتتــراوح قيمته بين الصفر (حجم العينة غير كاف ومن ثــم فإن التحليل العاملي ربما يكون غير ملائم ويصعب الحصول على عواملٍ ذات معنى) والواحد الصحيح (حجم العينة ملائم لتنفيذ التحليل العاملي ومن ثم الحصول على عوامل واضحة ذات معنى). وقد أشــار Field (2007, 640) إلى أن Kaiser (1997) أوصى بأنه إذا كانت قيمة KMO أكبر من أو تساوي ٠,٥ فإنه يمكن قبول حجم العينة لتنفيذ التحليــل العاملي، أما إذا كانت قيمــة KMO أقل من ذلك فإنه ينبغي زيادة حجم العينــة بجمع بيانات أكثر أو حذف بعض المتغيرات. وقد وضع Kaiser التصنيف التالي لقيم KMO وتفسيرها.

قيمة KMO	أقل من ٠,٥	٠,٥ - ٠,٧	٠,٧ - ٠,٨	٠,٨ - ٠,٩	٠,٩ أو أكبر
التفسير	حجم العينة غير كافٍ	حجم العينة متوسط	حجم العينة جيد	حجم العينة كبير	ممتاز

٤- أن تكون العلاقة بين المتغيرات خطية.

٥- أن يكون هناك ارتباط بين المتغيرات، والمقترح أن لا يقل الارتباط بين أي متغيرين عن ٠,٣٠.

٦- عدم وجود قيم شاذة مؤثرة.

٧- أن يكون توزيع المتغيرات محل الدراسة طبيعياً أو متماثلاً على الأقل.

خطوات تنفيذ التحليل العاملي الاستكشافي:

إن المخطط الانســيابي في الشــكل (١٠-٢) أدناه يوضح باختصار خطوات تنفيذ التحليل العاملي الاستكشافي في الواقع العملي التطبيقي.

شكل رقم (١٠-٢)
خطوات تنفيذ التحليل العاملي الاستكشافي
والسطور اللاحقة تقدم شرحاً كافياً لكل خطوة من تلك الخطوات

(1)

فحص وتهيئة البيانات (هل البيانات مؤهلة لإجراء التحليل العاملي)

- دراسة توزيع المتغيرات (طبيعي، متماثل أم ملتو بشدة، وجود القيم الشاذة)
- فحص معاملات الارتباط من خلال مصفوفة الارتباط، محددة مصفوفة الارتباط
- حجم العينة
- التحقق من صدق وثبات أداة القياس.
- التحقق من شرط التوزيع الطبيعي لمتغيرات

↓

(2)

استخلاص العوامل

هنالك عدة طرق من أكثرها استخداما وشيوعا بين الباحثين:
- تحليل المكونات الأساسية Principal Components Analysis (PCA)
- المحاور الأساسية Principal Axis Factoring (PAF)
- الاحتمال الأعظم Maximum Likelihood (ML)

↓

(3)

تدوير العوامل (المحاور)

هنالك نوعان من التدوير هما:
- التدوير المتعامد Orthogonal Rotation؛ ومن أكثر أساليبه استخداما التباين الأعظم Varimax
- التدوير المائل Oblique Rotation؛ ومن أكثر أساليبه استخداما البروماكس Promax



(4)

تفسير وتسمية (وصف) العوامل المستخلصة

ينبغي أن تكون العوامل المستخلصة ذات معنى منطقي أو متوافقة مع النظريات



(5)

إيجاد درجات العوامل

هنالك عدة طرق لإيجاد العوامل من أكثرها استخداما بين الباحثين
- طريقة جمع قياسات المتغيرات لكل عامل (إعادة ترميز المتغيرات أو العبارات العكسية)

أولاً - فحص البيانات والتحقق من ملاءمتها وجاهزيتها لتنفيذ التحليل العاملي:

قبـل تنفيـذ التحليل العاملـي واسـتخلاص العوامل يجـب التحقق مـن التالي (Field, 2007; Ho, 2005; Suhr, 2012):

١- موثوقية ودقة وكفاءة إجراءات القياس، والتحقق من صدق وثبات الاستبانة أو أداة القياس. ويقصد بالصدق هنا الصدق الظاهري وصدق المحتوى للاسـتبانة ويتم التحقق منه عن طريق الخبراء والمختصين في مجال الدراسـة. أما صدق البناء فيتم دراسته باسـتخدام التحليل العاملي، وقد سبق الإشارة إلى ذلك في الفصل الأول من الكتاب. أما بالنسبة للثبات، فسيتم الحديث عنه في نهاية هذا الفصل، علماً بأنه قد تم الإشارة إليه في الفصل الأول.

٢- حجم العينة. يجب أن يكون حجم العينة مناسباً وقد سبق تناوله سابقاً.

٣- التحقق من عدم وجود قيم شـاذة، حيث يتم اسـتخدام أساليب الإحصاء الوصفي الواردة في الفصل الرابع مثل التوزيع التكراري والقيم العظمى والصغرى للكشف عن القيم الشاذة ومعالجتها كما ورد في الفصل التاسع.

٤- وجود القيم المفقودة ومعالجتها وقد سبق الإشارة إلى ذلك في الفصل الخامس.

٥- فحص مصفوفة الارتباط. وللحصول على عوامل Factors أقل من عدد المتغيرات يجـب أن يكـون هناك ارتباط خطي بـين متغيرين على الأقل مـن المتغيرات في مصفوفـة الارتباط. ويمكن التحقق إحصائياً من ذلك من خلال فحص معاملات الارتبـاط الخطي في مصفوفة الارتباط. كما يمكـن التحقق من ذلك من خلال اختبار بارتليت Bartlett's Test for Sphericity والذي يختبر الفرضية:

H_0: مصفوفة الارتباط تسـاوي مصفوفة الوحدة (أي عدم وجود علاقات ارتباط خطي دالة إحصائياً بين المتغيرات).

مقابل:

H_1: مصفوفة الارتباط لا تسـاوي مصفوفة الوحـدة (أي وجود علاقات ارتباط خطي دالة إحصائياً بين المتغيرات).

وإحصـاءة اختبار بارتليت تتبع توزيع مربع كاي تقريباً، وهنا يتم مقارنة القيمة الاحتماليـة المحسـوبة من العينـة Sig. مع قيمة α المحددة سـلفاً ولتكن ٠,٠٥ فإذا كانت قيمة Sig. أقل من α فإننا نسـتنتج وجود علاقات ارتباط خطي دالة إحصائياً بين المتغيرات والعكس صحيح.

إن المتغير الذي ارتباطه ضعيف بالمتغيرات الأخرى يتم حذفه، كما أن المتغيرات ذات الارتباط العالي جداً $r > 0.9$ (مشكلة الازدواج الخطي العالي High Multicolinearity) ينبغي معالجتها إما بدمج المتغيرات أو بحذف أحد المتغيرات ذات الارتباط العالي جداً لأن الازدواج الخطي كما هو الحال في الانحدار الخطي المتعدد سيعيق تحديد مقدار المساهمة الجزئية للمتغيرات ذات الارتباط العالي في بناء أو تفسير العامل. كما يمكن كشف الازدواج الخطي العالي بين المتغيرات من خلال محددة مصفوفة الارتباط Correlation Matrix Determinant، فإذا كان محددة المصفوفة أكبر من ٠,٠٠٠٠١ فإنه يمكننا استنتاج عدم وجود مشكلة الازدواج الخطي العالي بين المتغيرات (Field, 2007).

٦- التحقق من شرط التوزيع الطبيعي (أو التماثل) لكل متغيرات الدراسة. وشرط التوزيع الطبيعي يعتبر هاماً في حال أردنا تعميم النتائج على المجتمع من بيانات العينة. وهذا الشرط ضروري فقط في حالة اختبار الفرضيات حول توزيع المتغيرات على العوامل كما في التحليل العاملي التوكيدي، أما في حالة التحليل العاملي الاستكشافي بطريقة المكونات الأساسية Principal Components يعتبر هذا الشرط غير ضروري (2007) Tabacknick & Fidell.

كما ينبغي إعادة ترميز الأسئلة أو العناصر أو البنود العكسية في الاستبانة أو أداة القياس في حال أردنا إيجاد درجات العوامل.

ثانياً - استخلاص العوامل Factors Extraction:

بعد التأكد من أن البيانات جاهزة لتنفيذ التحليل العاملي ننتقل إلى المرحلة الثانية والتي يتم فيها تحديد العوامل المناسبة والكافية لوصف البيانات، وذلك من خلال تقدير تشبعات العوامل على المتغيرات. ويتم استخلاص العوامل وفق ترتيب معين حسب نسبة تفسيرها للتباين المشترك في المتغيرات. حيث يتم أولاً استخلاص العامل الأكثر تفسيراً للتباين المشترك في البيانات، ثم يتم استخلاص العامل الذي يليه والذي يفسر أكبر نسبة من التباين المتبقي الذي لم يفسر بالعامل الأول، وهكذا حتى العامل الأخير. والجدير بالذكر أن عدد العوامل المستخلصة مبدئياً يساوي عدد المتغيرات. وحيث إن الهدف من التحليل العاملي تقليص عدد المتغيرات واختصارها في عدد أقل من العوامل، فإنه يتم الاقتصار على مجموعة العوامل التي تفسر الجزء الأكبر من التباين في المتغيرات والتي تكفي لفهم العلاقة بين المتغيرات، ومن ثم المفهوم محل الدراسة بالشكل الكافي والمفيد. ولاستخلاص العوامل فإن الباحث يقوم بما يلي:

١- تحديد طريقة استخلاص العوامل Extraction Method:

هناك طريقتان أساسيتان لاستخلاص العوامل هما:

- تحليل المكونات الأساسية Principal Components Analysis (PCA).
- التحليل العاملي المشترك Common Factor Analysis (CFA) ولهذه الطريقة عدة أساليب لاستخلاص العوامل، منها:

طريقة المحاور الأساسية Principal Axis Factoring (PAF)، والاحتمال الأعظم Maximum Likelihood (ML)، المربعات الصغرى غير الموزونة Unweighted Least Squares، المربعات الصغرى المعممة Generalized Least Squares، استخلاص العوامل ألفا Alpha Factoring، واستخلاص العوامل الصوري (التخيلي) Image Factoring. وتجدر الإشارة إلى أن SPSS يوفر تلك الأساليب.

وتختلف الطريقتان السابقتان PCA وCFA في كيفية تقدير شيوع المتغيرات Communalities المستخدمة في استخلاص العوامل أو المكونات بالإضافة إلى بعض الفروقات الرياضية والحسابية لكل منهما. وتعتبر طريقة المكونات الأساسية PCA، طريقة المحاور الأساسية PAF، والاحتمال الأعظم ML هي الأكثر استخداماً بين الباحثين (Thompson, 2004).

٢- تحديد عدد العوامل التي سيتم استخلاصها:

لقد سبق الإشارة إلى أن عدد العوامل المستخلصة مبدئياً يساوي عدد المتغيرات. وحيث إن الهدف من التحليل العاملي هو الاقتصار على عدد قليل من العوامل الممثلة لتلك المتغيرات، فإنه لابد من توافر معايير معينة تساعد الباحث على تحديد تلك العوامل. في أدبيات التحليل العاملي يتوفر العديد من المعايير أو الاختبارات لتحديد عدد العوامل المستخلصة، ومن أشهر تلك المعايير وأكثرها استخداماً في الواقع العملي ما يلي:

- الجذر الكامن Eigenvalue. أوصى كايزر Kaiser (1960) الاقتصار على العوامل التي لها جذور كامنة أكبر من أو تساوي الواحد الصحيح فقط. والجدير بالذكر أن معيار الجذر الكامن هو المعيار الافتراضي لتحديد عدد العوامل في برنامج SPSS.
- الرسم البياني Scree plot. وهو عبارة عن رسم بياني يوضح العلاقة بين قيمة الجذر الكامن للعامل (على المحور الرأسي) والعوامل المستخلصة المناظرة لها (على المحور الأفقي) كما في الشكل (١٠-٣) أدناه (Cattell, 1966). ويتضح من الرسم البياني أن العومل مرتبة تنازلياً حسب قيمة الجذر الكامن، فالعامل رقم ١ يقابل

الجذر الكامن الأكبر قيمة والعامل رقم ٢ يقابل الجذر الكامن الأقل قيمة من سابقه، وهكذا نستمر تنازلياً حتى العامل الأخير. وهنا يتم الاقتصار على العوامل التي تقع قبل المرفق أو نقطة الانحناء، أي أنه من الرسم البياني سنقتصر على اختيار العاملين الأولين ١ و٢.

شكل رقم (١٠-٣)

رسم الذراع للعلاقة بين قيم الجذور الكامنة والعوامل المناظرة لها

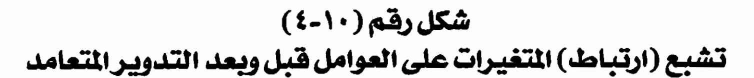


- نسبة التباين الكلي المفسر (المستخلص). وهو نسبة ما تفسره العوامل المستخلصة من التباين في البيانات (أو المتغيرات)، والحد الأدنى المقبول ٦٠٪ Hair, Anderson and Tatham (1998).
- شيوع المتغير على العامل. أن لا يقل عن ٠,٥٠ كما اقترح ذلك Larose.
- معيار التفسير Interpretability criterion. بعد اختيار عدد العوامل الممثلة للمتغيرات بالاعتماد على المعايير السابقة ينبغي أن تحقق التالي:
- القدرة على تسمية العامل بناء على المتغيرات المكونة له بالشكل المنطقي الذي يعكس معنى تلك المتغيرات مجتمعة.
- أن لا يقل تشبع المتغير على العامل عن القيمة ما بين ±٠,٣٠ و±٠,٤٠ حسب Tabacknick & Fidell بينما يرى Ho أن يكون ± ٠,٣٣، وأوصى Field باستخدام ±٠,٤٠. وأرى أن يبدأ الباحث بالقيمة ±٠,٤٠ فإذا توصل إلى عوامل تتفق مع المنطق أو النظرية، فهذا جيد وإلا فإنه ينتقل إلى الأقل منها مثلاً ± ٠,٣٣ أو ± ٠,٣٠.
- أن يكون ارتباط (تشبع المتغير على العامل) المتغير بعامل واحد فقط عالياً ومنخفضاً مع بقية العوامل.

وتجدر الإشارة إلى أن هذا المعيار بالإضافة إلى نسبة التباين المشترك في المتغيرات المفسرة بواسطة العوامل المستخلصة هو المحك الرئيسي في قبول الحل النهائي لنتائج التحليل العاملي الاستكشافي.

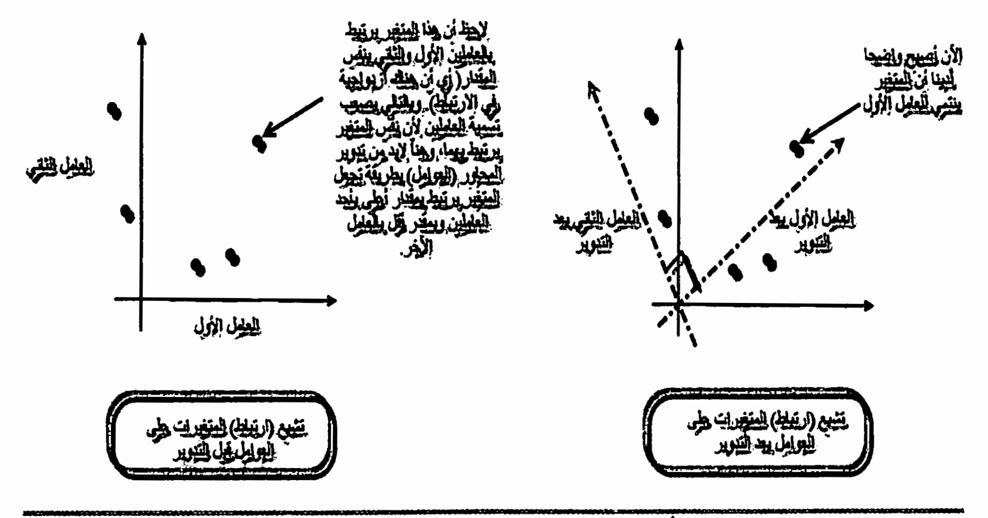
ثالثاً - تدوير العوامل المستخلصة Rotation of Extracted Factors:

فـــي الغالب فإن العوامل المســـتخلصة مبدئيـــاً يصعب تفســـيرها ويصعب إعطاء تســـميات منطقية لتلك العوامل، والســـبب يعود في ذلك إلى أنه فـــي المرحلة الأولى لاســـتخلاص العوامل يكون هناك متغيرات ترتبط بشكل عالٍ مع أكثر من عامل، ومن ثم يتولد لدينا مشـــكلة تتمثل في تحديد العوامل التي يرتبطُ بها ذلك المتغير بشـــكل معنوي، وهذه المشكلة تعرف بما يسمى «التشبع المتقاطع cross-loading» للمتغير على العوامل، ومن ثم يصعب إعطاء تســـمية منطقية لتلك العوامل. لذا نحتاج إلى تدوير العوامل (المحاور) بالدرجة التي تجعل كل متغير يتشبع (يرتبط بشكل كبير أو معنوي) مع عامل واحد فقط، في حين يكون تشـــبعه (ارتباطه) مع بقية العوامل ضعيفاً، وهذا يحقق لنا ما أسماه ثيرستون «البناء البسيط» (Field, 2007). والشكل (١٠-٤) يصور لنا بيانياً كيف أن هناك متغيراً متشـــبعاً (مرتبطاً) بنفس المقدار تقريباً على العاملين الأول والثاني في المرحلة المبدئية لاستخلاص العاملين مما يحدو بنا إلى القيام بتدوير المحاور بزاوية معينة لحل تلك المشكلة.



شكل رقم (١٠-٤)

تشبع (ارتباط) المتغيرات على العوامل قبل وبعد التدوير المتعامد



يوجد نوعان رئيسيان لتدوير العوامل وهما يختلفان حسب الزاوية التي بين العوامل بعد التدوير وهما:

- التدويـــر المتعامـــد Orthogonal Rotation: وفـــي هذا النوع مـــن التدوير يفترض الباحث أن العوامل مستقلة بعضها عن بعض، لذا يتم تدوير العوامل الأساسية بزاوية معينـــة مع الاحتفاظ بالتعامد (أي زاوية مقدارهـــا ٩٠٥) بين العوامل بعد التدوير، انظر الشـــكل (١٠-٤). ولهذا النوع من التدوير ثلاث طرق رئيسية هي: فاريماكس Varimax أي التباين الأعظم، كوارتيماكـــس Quartimax، وإكويماكس Equimax، ويعتبر أشهرها وأكثرها استخداماً بين الباحثين طريقة الفاريماكس كونها تعمل على فصل العوامل عن بعضها بشكل أوضح وأبسط في التكوين مع الحفاظ على التعامد بينها، وذلك على العكس من الطريقتين الأخريين (Ho, 2005).
- التدويـــر المائل Oblique Rotation: وفيه يعتقد الباحث أن العوامل مترابطة وغير مســـتقلة عن بعضها البعض، لذا يتم تدوير العوامل الأساسية بزاوية معينة مناسبة مـــع عدم الاحتفاظ بالتعامد بين العوامـــل (أي تكون الزاوية بين أي عاملين أقل من ٩٠٥) بعد التدوير، ولهذا النوع من التدوير عدة طرق منها طريقة أوبليمن Oblimin وطريقة بروماكس Promax، وتعتبر الأولى الأكثر استخداماً بين الباحثين.

رابعاً - تفسير أوتسمية أو وصف العوامل Interpreting Factors:

فـــي هذه المرحلة يتم تحديـــد المتغيرات المرتبطة بالعوامل المســـتخلصة من خلال حجم التشـــبعات (معاملات الارتباط بين المتغيرات والعوامل). فالمتغيرات التي يكون تشـــبع (ارتباط) العامل عليها أكبر من غيرها من المتغيرات الأخرى، تشير إلى أن تلك المتغيـــرات تنتمي إلى ذلك العامل. فعلى ســـبيل المثال، إذا كان لدينا ١٠ متغيرات وتم اســـتخلاص عاملين منها، فإنه ســـيتم توزيع تلك المتغيرات على العاملين حسب حجم التشـــبع (درجة الارتباط بين العامل والمتغير). فقد تكون درجة ارتباط المتغيرات من ١- ٦ بالعامل الأول أعلى من درجة ارتباط المتغيرات الأخرى ٧-١٠، وفي المقابل تكون درجة ارتباط المتغيرات ٧ إلى ١٠ بالعامل الثاني أعلى من درجة ارتباط المتغيرات ١-٦. وعليه يتم التوصل إلى استنتاج أن العامل الأول يتكون من المتغيرات ١-٦ والعامل الثاني يكون المتغيرات ٧-١٠. وبعد تحديد العوامل فإنه يتم تفســـيرها وتســـميتها، حيث يتم اشـــتقاق أو استنتاج اســـم العامل من المتغيرات التي يحتويها مجتمعة شريطة أن يتفق تفسير وتسمية أو وصف العوامل مع المنطق أو النظرية.

إضاءة إحصائية حول التحليل العاملي:

كما رأينا سابقاً فإن هناك العديد من الطرق لاستخلاص العوامل من المتغيرات، والعديد من المعايير المستخدمة في تحديد عدد العوامل المراد استخلاصها. كما أن هناك اختلافاً في الآراء ووجهات النظر بين الباحثين حول التطبيق العملي لتلك الطرق والمعايير والاختيار بينها. كل ذلك سيؤدي حتماً إلى حلول مختلفة. والمحك الرئيسي في التحليل العاملي هو الوصول إلى عدد قليل من العوامل يتفق تفسيرها ووصفها أو تسميتها مع المنطق أو النظرية حول المفهوم المراد قياسه بحيث:

- لا يقل الجذر الكامن للعامل عن ١.
- تفسر تلك العوامل نسبة كبيرة من التباين الكلي في البيانات لا يقل عن ٦٠٪ تقريباً.
- يكون ارتباط المتغيرات بالعوامل التي تنتمي إليها لا يقل عن ٠,٣٠.
- شيوع المتغير لا يقل عن ٠,٢٠، (Tabacknick & Fidell, 2007).

كيفية تنفيذ التحليل العاملي الاستكشافي باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سيتم توضيح كيفية تنفيذ التحليل العاملي الاستكشافي وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (١٠-١):

في بحث ميداني لدراسة «أسباب عدم الاستجابة في المسوح الميدانية: بحث ميداني على متدربي معهد الإدارة العامة»، قام الباحث بتصميم مقياس أو استبانة تتضمن ٣٠ عبارة أو بند أو سؤال أو متغير - انظر الملحق (٢) - يرى أنها تمثل في مجملها أبعاد أو جوانب الأسباب المحتملة لعدم الاستجابة في المسوح الميدانية (إسماعيل، ٢٠١١). وقد قام الباحث باستخدام مقياس ليكرت الخماسي من ١ إلى ٥ حيث تعني (٥) موافق بشدة، (٤) موافق، (٣) موافق إلى حد ما، (٢) غير موافق، (١) غير موافق بشدة، وذلك لتمثيل اتجاهات المستجيبين حول موضوع البحث.

ولتحقيق أهداف الدراسة، والتعرف على أهم أبعاد (عوامل) أسباب عدم الاستجابة في المسوح الميدانية من وجهة نظر المبحوثين، تم استخدام التحليل العاملي الاستكشافي بهدف فرز وتصنيف الـ (٣٠) عبارة (متغير) وتقليصها في عدد أقل من العوامل أو الأبعاد تفسر أكبر قدر ممكن من التباين الكلي في هذه المتغيرات.

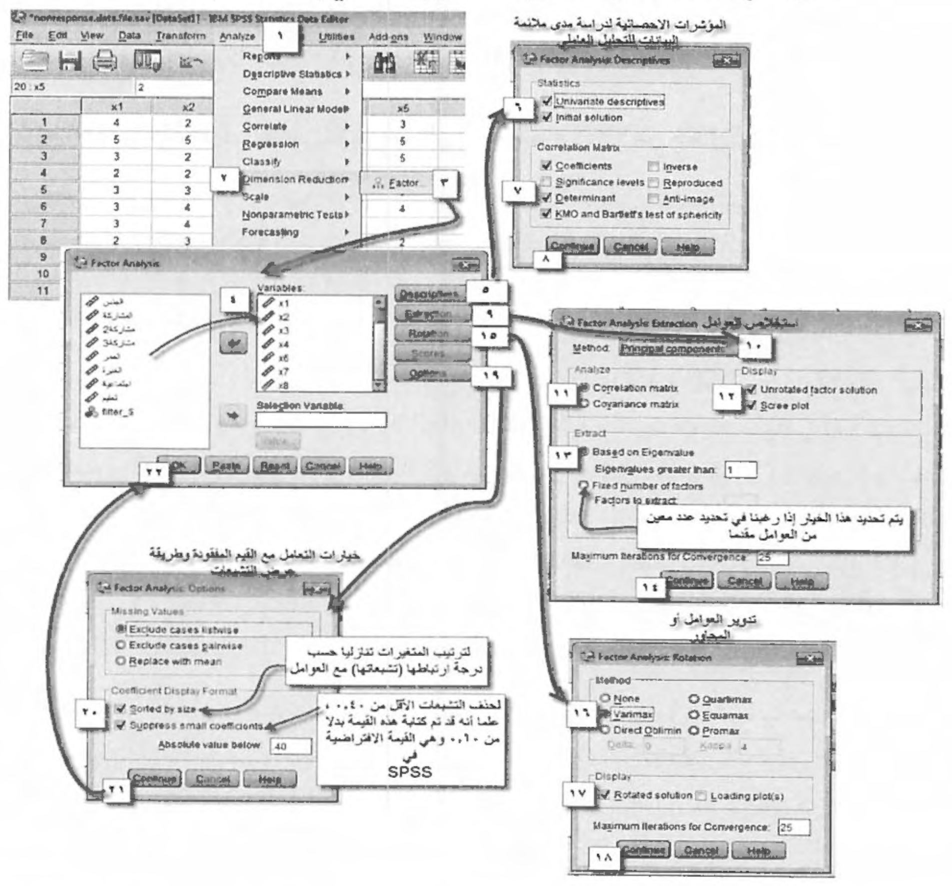
المطلوب:

إجراء التحليل العاملي باستخدام طريقة تحليل المكونات الأساسية PCA مع تدوير العوامل المســتخلصة (المحاور) بطريقة الفاريماكس (التباين الأعظم) Varimax، ومن ثم تفسير أو تسمية العوامل المستخلصة، ومن ثم توثيق النتائج وفقاً لأسلوب APA.

خطوات الحل:

لإجراء التحليل العاملي، قم أولاً بإدخال المتغيرات والبيانات بعد ترميزها إلى SPSS من خلال نافذتي Variable View وData View ومن ثم قم بحفظ الملف. بعد ذلك تتبع الخطوات حسب الأرقام التسلسلية من ١ إلى ٢٢ كما هو موضح بالشكل (١٠-٥) التالية:

شكل رقم (١٠-٥)
الخطوات التسلسلية لتنفيذ التحليل العاملي باستخدام SPSS



وبعد التنفيذ يتم الحصول على كم كبير من مخرجات التحليل والتي سنستخدمها لدراسة قابلية البيانات للتحليل العاملي ومن ثم استخلاص العوامل وتفسيرها.

في هذا التطبيق العملي تم تحليل بيانات ٧٦٢ مشاركاً تم قياس اتجاهاتهم نحو «أسباب عدم الاستجابة في المسوح» من خلال ٣٠ متغيراً، حيث تم بداية إلقاء نظرة استكشافية على توزيع المتغيرات وتبين أن توزيع جميع المتغيرات مقبول من حيث التماثل حيث لم يلاحظ وجود التواء شديد في توزيعاتها، كما اتضح أنه لا يوجود قيم شاذة تؤثر سلباً على التحليل. وينبغي الإشارة إلى أنه قد تم تنفيذ التحليل العاملي عدة مرات للتأكد من جاهزية البيانات للتحليل العاملي من خلال فحص مصفوفة الارتباط، وتحديد العوامل المنطقية المراد استخلاصها التي لها جذور كامنة Eigenvlaues لا تقل عن الواحد الصحيح، والتي تفسر مجتمعة ما لا يقل عن ٦٠٪ من التباين (Tatham, 1998) ومعاملات شيوع Communalities للمتغيرات على العوامل لا تقل عن ٠,٥٠ لكل متغير (Larose, 2006). وبناء عليه فقد تم حذف (٧) متغيرات مع الاحتفاظ بالمتبقي في التحليل النهائي للتحليل العاملي، والمتغيرات التي تم حذفها هي:

X٥: يحتوي الاستبيان على أسئلة كثيرة مفتوحة.

X١١: يحتوي الاستبيان على أسئلة إيحائية.

X١٢: يحتوي الاستبيان على أسئلة غير ضرورية.

X١٣: يحتوي الاستبيان على أسئلة افتراضية.

X١٥: خيارات الاجابات المحتملة للأسئلة غير كافية.

X٢٩: تصميم الاستبيان غير ملائم.

X٣٠: من أساسيات جهة العمل عدم المشاركة في أي استبيان.

والآن سنتعرف على خطوات إجراء التحليل العاملي للمتبقي من المتغيرات والبالغ عددها ٢٣ متغيراً.

أولاً - مدى قابلية وجاهزية البيانات للتحليل العاملي:

وقد تم فحص ما يلي:

- حجم العينة: حيث بلغ ٧٦٢ مشاركاً وهو أعلى بكثير من الحد الأدنى الموصى به لتنفيذ التحليل العاملي (٣٠٠ مشارك). كما يشير مؤشر KMO والذي بلغت قيمته ٠,٩٠٩، كما يتضح من الشكل (١٠-٧)، وهو أكبر من ٠,٩٠ بأن حجم العينة ملائم جداً.

- بفحص مصفوفة الارتباط الخطي اتضح أنه لا يوجد أي متغير جميع معاملات ارتباطه الخطي بأي متغير آخر تقل عن ٠,٣٠. كما أن اختبار بارتليت – انظر الشكل (١٠-٧) – لاختبار وجود علاقات ارتباط بين المتغيرات دالة إحصائياً من خلال الفرضيات التالية:

H_0: مصفوفة الارتباط تساوي مصفوفة الوحدة (أي عدم وجود علاقات ارتباط خطي بين المتغيرات).

مقابل

H_1: مصفوفة الارتباط لا تساوي مصفوفة الوحدة (أي وجود علاقات ارتباط خطي بين المتغيرات).

يؤكد وجود علاقات ارتباط خطي ذي دلالة إحصائية $(\chi^2_{0.05}(253) = 7968.74, p\text{ - }value < 0.001)$.

- عدم وجود مشكلة الازدواج الخطي حيث إنه لا يوجد أي متغيرين لهما معامل ارتباط أكبر من ٠,٩٠، وتؤكد محددة مصفوفة الارتباط ذلك، حيث إن قيمة المحددة والتي تساوي ٠,٠٠٠٠٢٥١٧ كما يتضح من الشكل (١٠-٦) أكبر من القيمة ٠,٠٠٠٠١.

شكل رقم (١٠-٦)

مقطع جزئي من مصفوفة الارتباط الخطي بين متغيرات الدراسة كما هو ظاهر في نافذة المخرجات

Correlation Matrix[a]

أسئلة الاستبيان تحتاج إلى تفكير عميق للإجابة عليها	يستغرق الاستبيان وقتاً طويلاً للإجابة على جميع أسئلته	الصياغة اللغوية لأسئلة الاستبيان غير سليمة	يحتوي الاستبيان على أسئلة ذات خصوصية	يحتوي الاستبيان على أسئلة مركبة (السؤال يتضمن أكثر من سؤال)	يحتوي الاستبيان على أسئلة متكررة	ليس لدي الوقت الكافي لتعبئة الاستبيان	عدم قناعتي بأهداف البحث موضوع الاستبيان
.355	.470	.290	.244	.297	.281	.360	.257
.346	.373	.395	.272	.362	.349	.233	.281
.335	.406	.430	.254	.273	.336	.239	.383
.345	.360	.417	.275	.362	.433	.209	.312
.436	.560	.403	.356	.405	.372	.409	.307
.509	.500	.476	.374	.473	.457	.319	.271
.525	.523	.518	.404	.463	.451	.312	.286
1.000	.626	.396	.373	.466	.389	.296	.252
.626	1.000	.431	.435	.477	.404	.467	.319
.396	.431	1.000	.407	.429	.457	.280	.320
.373	.435	.407	1.000	.461	.361	.273	.329
.466	.477	.429	.461	1.000	.556	.286	.295
.389	.404	.457	.361	.556	1.000	.281	.287

a. Determinant = 2.517E-005

قيمة محددة مصفوفة الارتباط وهي أقل من ٠,٠٠٠٠١

شكل رقم (١٠-٧)

نتائج اختبار KMO وبارتليت Bartlett

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.909
Bartlett's Test of Sphericity	Approx. Chi-Square	7968.735
	df	253
	Sig.	.000

ثانياً - استخلاص العوامل:

كمـا يتضح من الشـكل (١٠-٥) فـي الخطوات من ٩ إلى ١٣، فقد تم اسـتخدام طريقة تحليل المكونات الأساسـية PCA لاستخلاص العوامل، وبتنفيذ تلك الخطوات تم الحصول على معلومات حول الجذور الكامنة، الشيوع، نسبة التباين المفسر بالعوامل المسـتخلصة، والرسم البياني Scree plot والتي سنستخدمها في تحديد عدد العوامل المستخلصة.

- إن الشكل (١٠-٨) يوضح جزءاً من شيوع Communalities المتغيرات على العوامل، وحيث إن جميع الشيوع لا تقل عن ٠,٥٠ فإن هذا يقترح عدم حذف أي متغير من التحليل.

شكل رقم (١٠-٨)

مقطع جزئي من شيوع المتغيرات على العوامل كما هو في نافذة المخرجات

Communalities شيوع المتغيرات

	Initial	Extraction
الاستبيان طويل من حيث عدد الأسئلة المطلوب الإجابة عليها	1.000	.654
عدم وضوح الإرشادات اللازمة لتعبئة الاستبيان	1.000	.594
عدم وضوح الهدف من البحث	1.000	.586
أسئلة الاستبيان غير متسلسلة (ترتيبها غير منطقي)	1.000	.525
أسئلة الاستبيان طويلة	1.000	.613
يحتوي الاستبيان على مصطلحات غير مفهومة	1.000	.645
يحتوي الاستبيان على كلمات غير مفهومة	1.000	.655
أسئلة الاستبيان تحتاج إلى تفكير عميق للإجابة عليها	1.000	.563
يستغرق الاستبيان وقتاً طويلاً للإجابة على جميع أسئلته	1.000	.635
الصياغة اللغوية لأسئلة الاستبيان غير سليمة	1.000	.554
يحتوي الاستبيان على أسئلة ذات خصوصية	1.000	.512

من الشكل (١٠-٩) يتضح أن:

- عدد العوامل التي لها جذور كامنة لا تقل عن الواحد الصحيح يساوي (٥) عوامل.
- نســبة التباين الكلي المفســر بتلك العوامل الخمسة يساوي ٦٠,٩٦ ٪، وهو أكبر من الحد الأدنى الموصى به وهو ٦٠٪.

شكل رقم (١٠-٩)

التباين الكلي المفسر بواسطة العوامل المستخلصة

التباين الكلي المفسر بواسطة العوامل

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.069	35.084	35.084	8.069	35.084	35.084	3.748	16.297	16.297
2	2.316	10.070	45.154	2.316	10.070	45.154	2.885	12.544	28.841
3	1.466	6.375	51.529	1.466	6.375	51.529	2.762	12.007	40.849
4	1.154	5.018	56.547	1.154	5.018	56.547	2.683	11.666	52.515
5	1.015	4.413	60.960	1.015	4.413	60.960	1.942	8.446	60.960
6	.944	4.105	65.065						
7	.800	3.479	68.545						
8	.717	3.119	71.663						
9	.657	2.858	74.521						
10	.631	2.745	77.266						
11	.578	2.514	79.781						
12	.550	2.390	82.170						
13	.530	2.303	84.473						
14	.500	2.174	86.647						
15	.484	2.106	88.754						
16	.440	1.915	90.669						
17	.430	1.870	92.539						
18	.379	1.649	94.187						
19	.356	1.550	95.737						
20	.299	1.298	97.035						
21	.277	1.206	98.241						
22	.227	.986	99.227						
23	.178	.773	100.000						

Extraction Method: Principal Component Analysis.

العوامل

الجذور الكامنة

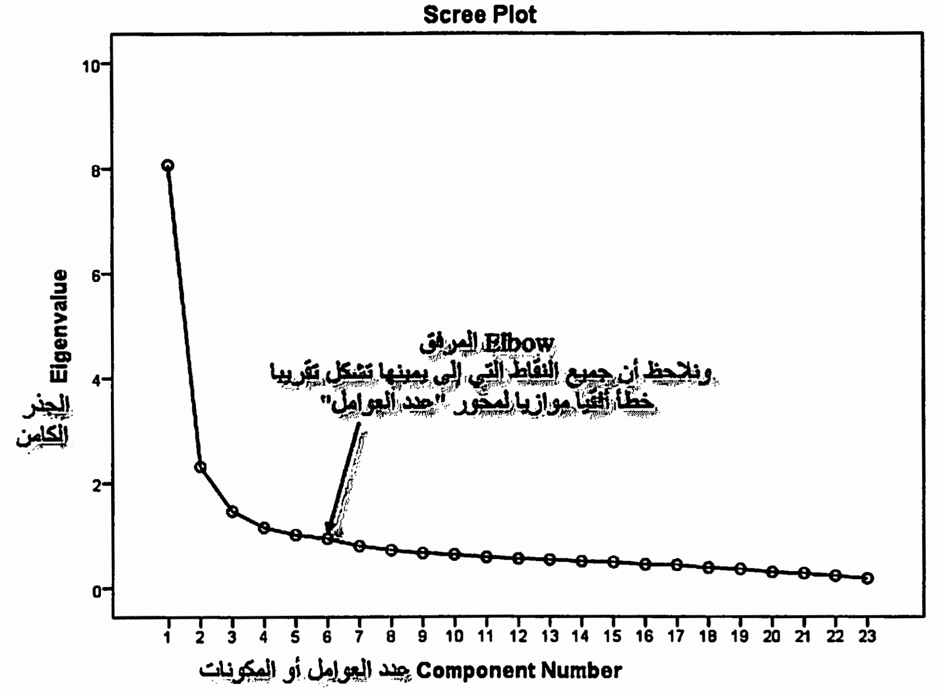
نسبة التباين الذي يفسره كل عامل من العوامل الخمسة المستخلصة في البيانات

نسبة التباين التراكمي للعوامل الخمسة ويتضح أن العوامل الخمسة مجتمعة تفسر ٦٠,٩٦% من التباين الكلي في البيانات

كما نلاحظ من الشــكل (١٠-١٠) أن الرسم البياني Scree Plot يشكل منعطفاً أو زاوية انحناء أو كما يســمى مرفقاً elbow عند العامــل رقم ٦ لذا يتم الاقتصار على العوامل التي قبله وهي الخمسة عوامل الأولى.

شكل رقم (١٠-١٠)

الرسم البياني Scree plot للعلاقة بين الجذور الكامنة وعدد العوامل كما هو ظاهر في نافذة المخرجات



إذن يتضح لنا من مما ســبق أن عدد العوامل المســتخلصة المقترحة والذي تحقق شروط الجذر الكامن، ونسبة التباين الكلي وشيوع المتغيرات يساوي (٥) عوامل.

ثالثاً - تدوير العوامل (المحاور) المستخلصة:

إن الخطــوات ١٥ إلى ١٨ في الشــكل (١٠-٥) تمكننا من تدوير العوامل (المحاور) بطريقــة الفاريماكس (التبايــن الأعظم) Varimax والتي تســاعد في الحصول على توزيــع أكثر وضوحاً للمتغيرات على العوامل المســتخلصة، كمــا أن الخطوات ١٩ إلى ٢١ في نفس الشــكل ستعمل على تحســين جودة عرض توزيع المتغيرات على العوامل،

وذلك من خلال ترتيب متغيرات كل عامل تنازلياً حسب حجم التشبعات Loadings مع استبعاد التشبعات التي تقل عن ٠,٤٠.

والشـكل رقم (١٠-١١) يعرض لنا العوامل المستخلصة بعد تدويرها باستخدام أسلوب الفاريماكس.

شكل رقم (١٠-١١)
العوامل المستخلصة بعد التدوير كما هو ظاهر في نافذة المخرجات

المتغيرات — مصفوفة العوامل بعد التدوير — العوامل

Rotated Component Matrix[a]

	Component 1	2	3	4	5
يحتوي الاستبيان على أسئلة مركبة (السؤال يتضمن أكثر من سؤال)	.718				
يحتوي الاستبيان على أسئلة متكررة	.698				
يحتوي الاستبيان على كلمات غير مفهومة	.665	.409			
يحتوي الاستبيان على مصطلحات غير مفهومة	.641	.432			
الصياغة اللغوية لأسئلة الاستبيان غير سليمة	.641				
يحتوي الاستبيان على أسئلة ذات خصوصية	.605				
الاستبيان طويل من حيث عدد الأسئلة المطلوب الإجابة عليها		.697			.430
عدم وضوح الإرشادات اللازمة لتعبئة الاستبيان		.689			
أسئلة الاستبيان تحتاج إلى تفكير عميق للإجابة عليها		.620			
عدم وضوح الهدف من البحث		.581			
أسئلة الاستبيان طويلة		.578			
أسئلة الاستبيان غير متسلسلة (ترتيبها غير منطقي)	.403	.495			
عدم ملائمة الوقت لتعبئة الاستبيان			.806		
عدم ملائمة المكان لتعبئة الاستبيان			.755		
ليس لدي الوقت الكافي لتعبئة الاستبيان			.745		
يستغرق الاستبيان وقتاً طويلاً للإجابة على جميع أسئلته	.429		.443		
لدي فكرة سالبة نحو موضوع البحث				.804	
لدي فكرة سالبة نحو المنظمة التي ينتمي إليها الباحث				.776	
لا أرى أي أهمية أو مردود للبحث العلمي عموماً				.696	
عدم قناعتي بأهداف البحث موضوع الاستبيان			.467	.530	
تتطلب الإجابة على أسئلة الاستبيان خلفية كبيرة بالموضوع					.726
أرى أن اختياري غير مناسب للإجابة على أسئلة الاستبيان					.649
أخشى من استخدام البيانات التي أدلي بها لغير أغراض البحث				.525	.545

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 8 iterations.

رابعاً - تفسير أو تسمية العوامل:

كما نلاحظ من الشكل (١٠-١١) أعلاه، فإنه تم استخلاص خمسة عوامل لأسباب عدم الاستجابة في المسوح يتشبع (يرتبط) كل منها على عدد من المتغيرات لا يقل عن ثلاثة، وتفسر في مجملها ما نسبته ٩٦, ٦٠٪ من التباين الكلي في البيانات، ويمكن تسمية هذه العوامل وفقاً لمضمونها بوجه عام على النحو التالي:

العامل الأول: ضعف صياغة الأسئلة:

احتوى هذا العامل على (٦) متغيرات أو عبارات تقيس في مجملها «ضعف صياغة الأسئلة» وتأتي هذه العوامل مرتبة حسب درجة تشبعاتها على العامل كالتالي:

- يحتوي الاستبيان على أسئلة مركبة، يحتوي الاستبيان على أسئلة متكررة، يحتوي الاستبيان على كلمات غير مفهومة، يحتوي الاستبيان على مصطلحات غير مفهومة، الصياغة اللغوية لأسئلة الاستبيان غير سليمة، يحتوي الاستبيان على أسئلة ذات خصوصية.

العامل الثاني: طول الاستبيان وعدم وضوح الارشادات والأسئلة:

احتوى هذا العامل (٦) متغيرات أو عبارات تقيس في مجملها «طول الاستبيان وعدم وضوح الارشادات والأسئلة» وتأتي هذه العوامل مرتبة حسب درجة تشبعاتها على العامل كالتالي:

- الاستبيان طويل من حيث عدد الأسئلة، عدم وضوح الإرشادات اللازمة لتعبئة الاستبيان، أسئلة الاستبيان تحتاج إلى تفكير عميق للإجابة عنها، عدم وضوح الهدف من البحث، أسئلة الاستبيان طويلة، أسئلة الاستبيان غير متسلسلة.

العامل الثالث: عدم ملاءمة وقت ومكان تنفيذ الاستبيان:

احتوى هذا العامل (٤) متغيرات أو عبارات تقيس في مجملها «عدم ملاءمة وقت ومكان تنفيذ الاستبيان» وتأتي هذه العوامل مرتبة حسب درجة تشبعاتها على العامل كالتالي:

- عدم ملاءمة الوقت لتعبئة الاستبيان، عدم ملاءمة المكان لتعبئة الاستبيان، ليس لدي الوقت الكافي لتعبئة الاستبيان، يستغرق الاستبيان وقتاً طويلاً للإجابة عن جميع الأسئلة.

العامل الرابع: النظرة السلبية اتجاه البحث العلمي:

تضمن هذا العامل (٤) متغيرات أو عبارات تقيس في مجملها «النظرة السلبية اتجاه البحث العلمي» وتأتي هذه العوامل مرتبة حسب درجة تشبعاتها على العامل كالتالي:

- لدي فكرة سالبة نحو موضوع البحث، لدي فكرة سالبة نحو المنظمة التي ينتمي إليها الباحث، لا أرى أهمية أو مردود للبحث العلمي عموماً، عدم قناعتي بأهداف البحث موضوع الاستبيان.

العامل الخامس: الخشية من عواقب المشاركة في المسوح:

تضمن هذا العامل (٣) متغيرات أو عبارات تقيس في مجملها «الخشية من المشاركة في الدراسات المسحية» وتأتي هذه العوامل مرتبة حسب درجة تشبعاتها على العامل كالتالي:

- تتطلب الإجابة عن أسئلة الاستبيان خلفية كبيرة بالموضوع، أرى أن اختياري غير مناسب للإجابة عن أسئلة الاستبيان، أخشى من استخدام البيانات التي أدلى بها لغير أغراض البحث.

خامساً - ملخص النتائج وتوثيقها:

وبعد تنفيذ التحليل العاملي وتفسير مخرجات التحليل والوصول إلى العوامل النهائية لأسباب عدم الاستجابة في المسوح وتسميتها، سنقوم الآن بتلخيص كل ما سبق مناقشته في هذا التطبيق استعداداً لتوثيق نتائج التحليل العاملي وفقاً لأسلوب APA.

توثيق النتائج:

تم تنفيذ التحليل العاملي الاستكشافي باستخدام طريقة تحليل المكونات الأساسية على بيانات ٧٦٢ مشاركاً ساهموا في تعبئة استبيان يتضمن ٣٠ بنداً أو متغيراً تقيس أسباب عدم الاستجابة في المسوح الميدانية، وقد تم حذف سبعة متغيرات معاملات شيوعها أقل من ٠,٥٠ حسب Larose (2006) قبل اعتبار التحليل النهائي للتحليل العاملي. وللتأكد من قابلية البيانات لإجراء التحليل العاملي، تم فحص مصفوفة الارتباط للمتغيرات المتبقية حيث تبين أن جميع المتغيرات لها معامل ارتباط خطي واحد على الأقل أكبر من ٠,٣٠.، وأكد اختبار بارتلت (Bartlett's Test of Sphericity) وجود علاقات ارتباط خطي دالة إحصائياً

بين المتغيرات ($\chi^2_{0.05}(253) = 7968.74, p\text{ - }value < 0.001$) مما يدل على قابلية البيانات لتكوين العوامل. كما أن قيمة مقياس Kaiser-Meyer-Olkin (KMO) لمدى ملاءمة حجم العينة للتحليل العاملي تساوي ٠,٩١ مما يعني أن حجم العينة ملائم جداً حسب تصنيف (Kaiser,1997).

لقد أسفر التحليل العاملي عن خمسة عوامل لها جذور كامنة أكبر من الواحد، تفسر مجتمعة ٦٠,٩٦٪ من التباين الكلي في البيانات وهو أكبر من الحد الأدنى الموصى به ٦٠٪ (Haire, et.al.,1998)، حيث يفسر كل منها ١٦,٣٠٪، ١٢,٥٤٪، ١٢,٠٠٪، ١١,٦٧٪، ٨,٤٤٪ من التباين الكلي على التوالي. كما أن معاملات الشيوع للمتغيرات على العوامل لا يقل أي منها عن (Larose, 2006) ٠,٥٠.. ويؤيد الرسم البياني Scree Plot استخلاص خمسة عوامل (Cattell,1966).

ولتفسير العوامل الخمسة وتسميتها، تم تدوير العوامل بأسلوب التباين الأعظم (الفاريماكس) Varimax والتي تعكس تحديداً واضحاً وبسيطاً لبنية العوامل (Ho, 2005). إن تفسير العوامل الخمسة وتوزيع المتغيرات على العوامل كان واضحاً ومنطقياً حيث تشبعت البنود التي تقيس ضعف صياغة الأسئلة بدرجة كبيرة على العامل الأول، وطول الاستبيان وعدم وضوح الإرشادات والأسئلة على العامل الثاني، وعدم ملاءمة وقت ومكان تنفيذ المسوح على العامل الثالث، النظرة السلبية تجاه البحث العلمي على العامل الرابع، والخشية من المشاركة في المسوح على العامل الخامس. والجدول (١٠-١) يعرض تشبعات المتغيرات على العوامل المستخلصة وشيوعها بعد تدوير العوامل.

جدول رقم (١٠-١)

نتائج التحليل العاملي لمقياس أسباب عدم الاستجابة في المسوح الميدانية

البند	تشبعات العوامل بعد التدوير Rotated Factor Loadings					الشيوع
	العامل ١	العامل ٢	العامل ٣	العامل ٤	العامل ٥	
يحتوي الاستبيان على أسئلة مركبة	٠,٧١٨					٠,٥٩٧
يحتوي الاستبيان على أسئلة متكررة	٠,٦٩٨					٠,٥٥٧
يحتوي الاستبيان على كلمات غير مفهومة	٠,٦٦٥					٠,٦٥٥
يحتوي الاستبيان على مصطلحات غير مفهومة	٠,٦٤١					٠,٦٤٥
الصياغة اللغوية لأسئلة الاستبيان غير سليمة	٠,٦٤١					٠,٥٥٤

تابع - جدول رقم (١٠-١).

البند	تشبعات العوامل بعد التدوير Rotated Factor Loadings					الشيوع
	العامل ١	العامل ٢	العامل ٣	العامل ٤	العامل ٥	
يحتوي الاستبيان على أسئلة ذات خصوصية	٠,٦٠٥					٠,٥١٢
الاستبيان طويل من حيث عدد الأسئلة		٠,٦٩٧				٠,٦٥٤
عدم وضوح الإرشادات اللازمة لتعبئة الاستبيان		٠,٦٨٩				٠,٥٩٤
أسئلة الاستبيان تحتاج إلى تفكير عميق للإجابة عنها		٠,٦٢٠				٠,٥٦٣
عدم وضوح الهدف من البحث		٠,٥٨١				٠,٥٨٦
أسئلة الاستبيان طويلة		٠,٥٧٨				٠,٦١٣
أسئلة الاستبيان غير متسلسلة		٠,٤٩٥				٠,٥٢٥
عدم ملاءمة الوقت لتعبئة الاستبيان			٠,٨٠٦			٠,٧٦٤
عدم ملاءمة المكان لتعبئة الاستبيان			٠,٧٥٥			٠,٦٩٨
ليس لدي الوقت الكافي لتعبئة الاستبيان			٠,٧٤٥			٠,٦٣٥
يستغرق الاستبيان وقتاً طويلاً للإجابة عن جميع الأسئلة			٠,٤٤٣			٠,٦٣٥
لدي فكرة سالبة نحو موضوع البحث				٠,٨٠٤		٠,٧٠٢
لدي فكرة سالبة نحو المنظمة التي ينتمي إليها الباحث				٠,٧٧٦		٠,٦٦٣
لا أرى أهمية أو مردوداً للبحث العلمي عموماً				٠,٦٩٦		٠,٥٥٠
عدم قناعتي بأهداف البحث موضوع الاستبيان				٠,٥٣٠		٠,٥٨١
تتطلب الإجابة عن أسئلة الاستبيان خلفية كبيرة بالموضوع					٠,٧٢٦	٠,٦٠٩
أرى أن اختياري غير مناسب للإجابة عن أسئلة الاستبيان					٠,٦٤٩	٠,٥٣٧
أخشى من استخدام البيانات التي أدلى بها لغير أغراض البحث					٠,٥٤٥	٠,٥٩١

الثبات Reliability:

لقد سبقت الإشارة في الفصل الأول إلى أنه يجب أن تكون أداة القياس صادقة وثابتة. حيث يمكن التحقق من صدق أداة القياس (صدق البناء) إحصائياً باستخدام التحليل العاملي، وذلك من خلال توزيع بنود أو عبارات الأداة على مجموعة من المكونات الأساسية أو العوامل. أما الثبات فإنه يمكن التحقق منه إحصائياً من خلال ما يسمى بمعاملات الثبات.

إن الثبات يقصد به إلى أي درجة يعطي المقياس قراءات متقاربة عند تطبيقه في كل مرة. فالأداة المتذبذبة التي تعطي نتائج متفاوتة عند تطبيقها أكثر من مرة تكون مدعاة للقلق وعدم الثقة في نتائجها ومن ثم تعد هدراً للجهد والوقت والمال. ويمكن حساب الثبات كمياً (ويطلق عليها معاملات الثبات) باستخدام الأساليب الإحصائية، وتتراوح قيمته بين الصفر (أداة منعدمة الثبات) والواحد الصحيح (أداة تامة الثبات). ويمكن تقسيم معاملات الثبات إلى ثلاثة أقسام حسب طريقة حساب معامل الثبات كالتالي Huck (2012):

١- **معاملات الاستقرار Coefficient of Stability**: ويعني ثبات المقياس خلال الزمن، ومن أشهر مقاييسه ما يسمى بمعامل «الاختبار- إعادة الاختبار» Test-Retest Reliability Coefficient وفيه يتم تطبيق الاختبار أو المقياس على نفس المجموعة من الأفراد أو الأشياء في فترتين زمنيتين مختلفتين، ومن ثم يتم حساب معامل الثبات للمقياس من خلال معامل الارتباط الخطي بين نتائج المقياس التي تم الحصول عليها في فترتي تطبيق المقياس.

٢- **معاملات التكافؤ أو النماذج البديلة Alternate-Forms Coefficients**: وفي هذه الطريقة يتم تصميم نموذجين متكافئين من المقياس ومن ثم تطبيقه على نفس المجموعة من الأفراد مع ترك فترة زمنية وجيزة بين التطبيق الأول والثاني للمقياس. ومن ثم يتم حساب معامل الثبات للمقياس من خلال معامل الارتباط الخطي بين نتائج المقياس التي تم الحصول عليها من التطبيق الأول والثاني للمقياس.

٣- **معاملات الثبات الداخلي Internal Consistency Reliability**: وفي هذه الطريقة يتم تطبيق المقياس مرة واحدة فقط على مجموعة من الأفراد. ويتم حساب معامل الثبات في هذه الحالة بإيجاد معامل الارتباط بين عناصر المقياس بعد تجزئتها بطريقة معينة إلى جزئين، ولهذا سميت هذه الطريقة بالثبات

الداخلي. ويوجد عدة طرق لحساب معامل الثبات الداخلي حسب طريقة تجزئة عناصر المقياس منها: معاملات التجزئة النصفية Split-half، معاملات كودر - ريتشردسون (Kuder-Richardson 20 (K-R 20 - وهي خاصة بالاستبانات ذات العبارات أو البنود الثنائية، أي التي تكون قيمها المحتملة قيمتين فقط (نعم أو لا، ٠ أو١) -، ومعامل ألفا كرونباخ Cronbach's Alpha. ويعد معامل ألفا كرونباخ من أكثر مقاييس الثبات الداخلي استخداماً في الواقع العملي ويتم حسابه وفق الصيغة الرياضية التالية:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma^2_{Y_i}}{\sigma^2_Y}\right)$$

حيث k تمثل عدد عناصر (بنود أو أسئلة) الاستبانة، σ^2_Y تمثل تباين بيانات الأفراد للعنصر أو البند i وσ^2_Y تمثل التباين الكلي لبيانات الأفراد لكل عناصر أو بنود المقياس.

وتتراوح قيمة ألفا بين الصفر والواحد. وبوجه عام يعد المقياس ذا ثبات منخفض إذا كانت ألفا أقل من ٠,٤٠ وذا ثبات متوسط إذا كانت قيمة ألفا تقع في الفترة بين ٠,٤٠ و٠,٧٠، وذا ثبات مرتفع إذا كانت ألفا ٠,٧٠ فأعلى. فإذا كان المقياس ذا ثبات عالٍ، فإن هذا يعني أن بنود المقياس ثابتة Reliable، ومن ثم المقياس ثابت، أما إذا كان الثبات منخفضاً فإن ذلك يعني أن هناك واحداً على الأقل من بنود أو عبارات المقياس غير ثابت، ومن ثم ينبغي تحديدها ومعالجة ذلك كأن يتم حذف البند أو العبارة من التحليل.

كيفية تحليل الثبات باستخدام SPSS من خلال تطبيق عملي تفاعلي:

فيما يلي سيتم توضيح كيفية تحليل الثبات وتفسير النتائج باستخدام برنامج SPSS من خلال التطبيق العملي التفاعلي التالي.

تطبيق عملي تفاعلي (١٠-٢):

في التطبيق العملي التفاعلي (١٠-١) السابق تم استخدام التحليل العاملي الاستكشافي، وذلك للتعرف على أهم أبعاد (عوامل) أسباب عدم الاستجابة في المسوح الميدانية من وجهة نظر المبحوثين، وفرز وتصنيف الـ (٣٠) عبارة أو بند (متغير) وتقليصها في عدد أقل من العوامل أو الأبعاد تفسر أكبر قدر ممكن من التباين الكلي

في هذه المتغيرات. وقد أسـفر التحليل العاملي عن تحديد واستخلاص خمسة عوامل من أصل (٢٣) متغيراً بعد اسـتبعاد سبعة بنود أومتغيرات. والجدول (١٠-١) يوضح توزيع المتغيرات أو البنود على العوامل الخمسة.

المطلوب: اختبار ثبات الاسـتبانة من خلال حساب معامل ألفا كرونباخ لكل عامل، ومن ثم توثيق النتائج وفقاً لأسلوب APA.

خطوات الحل:

أولاً - إدخال المتغيرات والبيانات ومن ثم تنفيذ التحليل:

قبل البدء في التحليل، ينبغي التأكد من عدم وجود بنود أو عبارات عكسـية لأنها تؤثـر على قيمة معامـل ألفا كرونباخ، وفي هذا التطبيق لا يوجـد لدينا بند أو عبارة عكسية. وفي حال وجوده ينبغي إعادة ترميزه.

وتجدر الإشـارة إلى أننا سنقوم بحساب معامل الثبات ألفا كرونباخ لكل عامل من العوامل الخمسة السابقة على حدة. وسنكتفي هنا بعرض كيفية حساب معامل الثبات للعامل الأول «ضعف صياغة الأسـئلة» والذي يتكون من سـتة بنود باستخدام SPSS ومن ثم يمكن حساب معامل الثبات لبقية العوامل بنفس الطريقة.

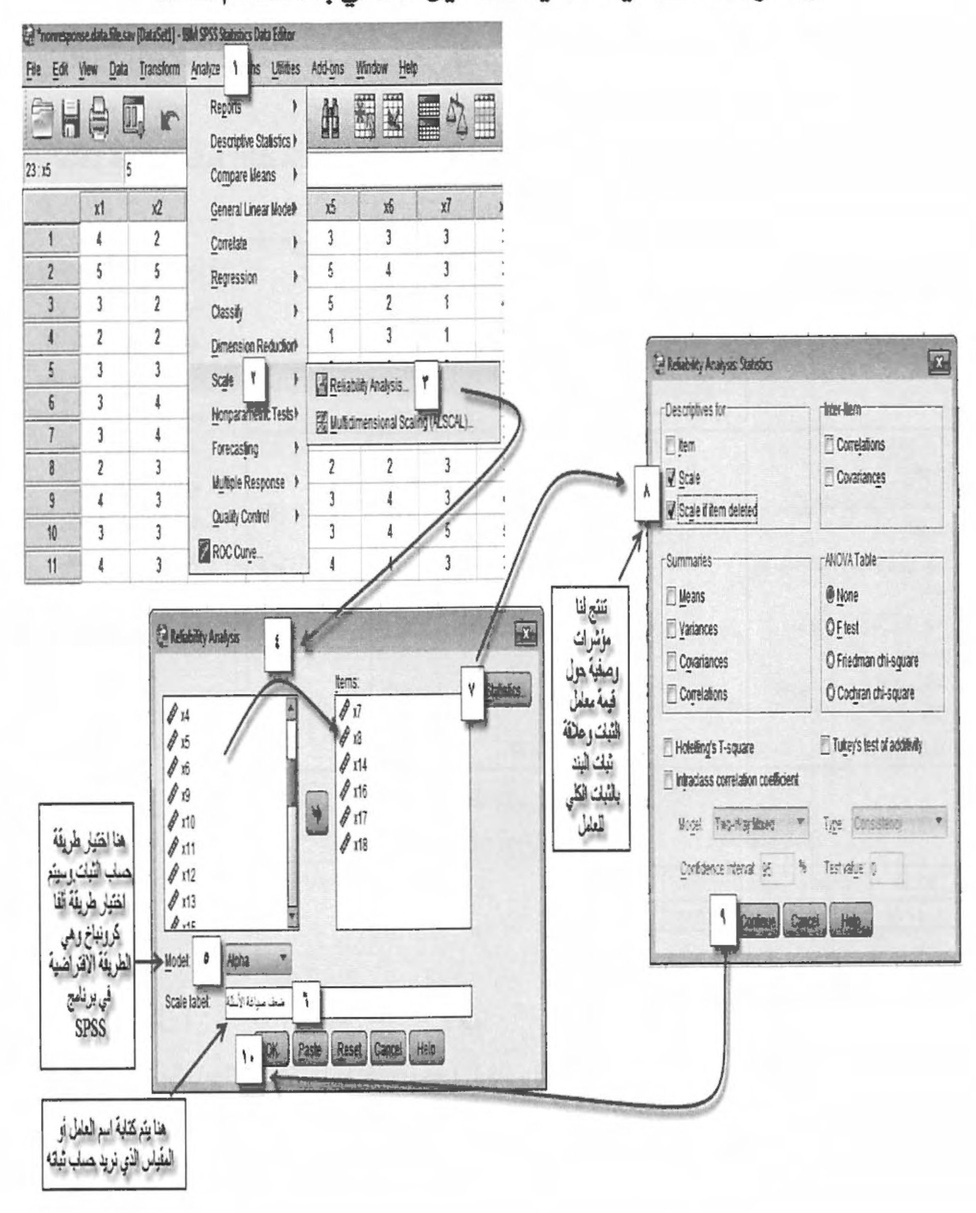
تحليل معامل الثبات ألفا كرونباخ للعامل «ضعف صياغة الأسئلة»:

في الواقع العملي يتم عادة تحليل الثبات بالاقتران مع التحليل العاملي، لذا سـيتم استخدام ملف SPSS للبيانات الذي تم إنشاؤه في التطبيق العملي التفاعلي (١٠-١).

لإجراء تحليل الثبات، قم بتتبع الخطوات حسـب الأرقام التسلسـلية من ١ إلى ١٠ كما هو موضح بالشكل (١٠-١٢) التالية:

شكل رقم (١٠-١٢)

الخطوات التسلسلية لتنفيذ التحليل العاملي باستخدام SPSS



ثانياً - قراءة نتائج التحليل وتفسيرها:

وبعد التنفيذ يتم الحصول على مجموعة من الجداول كما هو يوضح الشكل (١٠-١٣) في نافذة المخرجات يتم استخدامها في تحليل الثبات للعامل.

شكل رقم (١٠-١٣)

مخرجات تحليل الثبات (معامل ألفا كرونباخ) كما هي ظاهرة في نافذة المخرجات



Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
يحتوي الاستبيان على مصطلحات غير مفهومة	14.53	17.390	.678	.793
يحتوي الاستبيان على كلمات غير مفهومة	14.70	17.261	.696	.789
الصياغة اللغوية لأسئلة الاستبيان غير سليمة	14.92	18.526	.578	.814
يحتوي الاستبيان على أسئلة ذات خصوصية	14.83	19.100	.508	.827
يحتوي الاستبيان على أسئلة مركبة (السؤال يتضمن أكثر من سؤال)	14.55	18.372	.607	.808
يحتوي الاستبيان على أسئلة متكررة	14.69	18.409	.581	.813

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
17.64	25.335	5.033	6

من الشكل (١٠-١٣) يتضح لنا ما يلي:

- أن قيمة معامل الثبات (ألفا كرونباخ) الكلي للعامل «ضعف صياغة الأسئلة» والمكون من ستة بنود تساوي ٠,٨٣٤، وهذا يعني أن ثبات العامل عالٍ.
- في العمود المعنون بـ "Corrected Item-Total Correlation" نلاحظ أن ارتباط كل بنـد أو متغير ببقية البنـود الأخرى لا يقل عن ٠,٤٠، وهذا يقترح ألا يتم حذف أي من تلك البنود.

- في العمود المعنون بـ "Cronbach's Alpha if Item Deleted" والذي يعني «قيمة معامل ألفا كرونباخ للعامل بعد حذف البند المعني» ويتم اســتخدامه لاستكشاف ما إذا كان حذف البند سيؤدي إلى تحسين درجة الثبات للعامل حيث يتم مقارنة القيم في هذا العمود مع قيمة معامل ألفا كرونباخ الكلية للعامل المكون من كل البنود (ستة بنــود في مثالنا هذا)، فإذا كانت قيمة ألفا كرونباخ للعامل بعد حذف البند أكبر من قيمــة معامل ألفا كرونبــاخ الكلية للعامل بدون حذف أي بند، فإن هذا يقترح حذف ذلــك البنــد. فمثلاً نلاحظ أن قيمة معامل ألفا كرونبــاخ بعد حذف البند «يحتوي الاســتبيان على مصطلحات غير مفهومة» تســاوي ٠,٧٩٣، وهي أقل من قيمة ألفا كرونباخ الكلية (للستة بنود مجتمعة) والتي تساوي ٠,٨٣٤، وهذا يقترح عدم حذف البنــد. وبالنظر إلى بقية البنود فإننا نلاحظ أن معاملات ألفا كرونباخ لها أقل من ٠,٨٣٤، وهذا يعني أن حذف أي من تلك البنود لن يؤدي إلى تحســين درجة الثبات الكلية لبنود العامل.

نســتنتج مما سبق أنه لن يتم حذف أي من البنود المكونة للعامل، ومن ثم فإن قيمة معامل ألفا كرونباخ لهذا العامل والتي سيتم توثيقها في نتائج البحث هي ٠,٨٣٤.

وباتباع نفس الخطوات الســابقة لتحليل الثبات للعوامل الأخرى باستخدام SPSS فإن قيم معاملات ألفا كرونباخ للعوامل هي:

رقم العامل	اسم أو وصف العامل	عدد بنوده	قيمة ألفا كرونباخ
١	ضعف صياغة الأسئلة	٦	٠,٨٣
٢	طول الاستبيان وعدم وضوح الإرشادات والأسئلة	٦	٠,٨٠
٣	عدم ملاءمة وقت ومكان تنفيذ الاستبيان	٤	٠,٨١
٤	النظرة السلبية اتجاه البحث العلمي	٤	٠,٧٨
٥	الخشية من عواقب المشاركة في المسوح	٣	٠,٧١

ثالثاً - ملخص النتائج وتوثيقها:

للتحقق من ثبات الاستبيان «أسباب عدم الاستجابة في المسوح»، تم فحص الثبات الداخلــي لــكل عامل (بعد) من العوامل الخمســة المكونة لبنود الاســتبانة (٢٣ بند)، وذلك باســتخدام معامل ألفا كرونباخ. وكانت قيم ألفــا كرونباخ مرتفعة، حيث بلغت

٠,٨٣ للعامل «ضعف صياغة الأسئلة» (٦ بنود)، و٠,٨٠ للعامل «طول الاستبيان وعدم وضوح الإرشادات والأسئلة» (٦ بنود)، و٠,٨١ للعامل «عدم ملاءمة وقت ومكان تنفيذ الاستبيان» (٤ بنود)، و٠,٧٨ للعامل «النظرة السلبية اتجاه البحث العلمي» (٤ بنود)، و٠,٧١ للعامل «الخشية من عواقب المشاركة في المسوح» (٣ بنود). كما اتضح من نتائج التحليل أنه لم يكن هناك زيادة كبيرة في قيم معاملات ألفا كرونباخ لأي عامل من العوامل الخمسة بحذف بند أو أكثر.

ملاحظة:

عند استخدام التحليل العاملي وألفا كرونباخ لاختبار صدق البناء والثبات الداخلي للاستبانة على التوالي، فإنه يتم أولاً توثيق نتائج التحليل العاملي يليه مباشرة توثيق نتائج تحليل الثبات الداخلي.

إضاءات إحصائية حول الثبات الداخلي:

- إذا تم حذف أحد البنود لتحسين درجة الثبات الداخلي للعامل (البعد) أو الاستبانة فإنه ينبغي إعادة تنفيذ التحليل العاملي للتأكد من أن تركيبة العوامل لم تتغير.
- كما يتضح من الصيغة الرياضية لمعامل ألفا كرونباخ، فإنه من المتوقع أن تزيد قيمة معامل ألفا بزيادة عدد البنود. فالقيمة العالية لمعامل ألفا كرونباخ قد لا تكون بسبب أن الثبات الداخلي لبنود المقياس عال، وإنما قد يكون بسبب العدد الكبير للبنود الذي يحتويها المقياس. فقد يكون هناك بنود زائدة عن الحاجة أو مكررة Redundant لن تعطي أي معلومات إضافية، ومن ثم فإن وجودها في المقياس سيضخم قيمة ألفا كرونباخ موحياً بأن الثبات الداخلي عالٍ، وهو قد يكون في الحقيقة عكس ذلك.

قائمة المراجع:

أولاً - المراجع العربية:

١- أبوشعر، عبدالرزاق أمين (١٩٩٧). **العينات وتطبيقاتها في البحوث الاجتماعية**. المملكة العربية السعودية، الرياض: معهد الإدارة العامة.

٢- إسماعيل، محمد عبدالرحمن (٢٠١١). أسباب عدم الاستجابة في المسوح الميدانية: **بحث مسحي على متدربي معهد الإدارة العامة**. معهد الإدارة العامة، ٤(٥١)، ٥١٧-٥٧٣.

٣- العساف، أحمد عارف والوادي، محمود (٢٠١١). **منهجية البحث في العلوم الاجتماعية والإدارية (المفاهيم والأدوات)**. الأردن، عمان: دار صفاء للنشر والتوزيع.

٤- القحطاني، سعد سعيد (١٤٣٢). **الأساليب الإحصائية في الدراسات الميدانية (حقيبة تدريبية)**. المملكة العربية السعودية، الرياض: معهد الإدارة العامة.

٥- النجار، فايز جمعة، النجار، نبيل جمعة و الزعبي، ماجد راضي (٢٠١٠). **أساليب البحث العلمي: منظور تطبيقي (الطبعة الثانية)**. الأدرن، عمان: دار الحامد للنشر والتوزيع.

٦- بري، عدنان ماجد، هندي، محمود محمد وراضي، الحسيني عبدالبر (١٩٩٨). **أساسيات طرق التحليل الإحصائي**. المملكة العربية السعودية، الرياض: جامعة الملك سعود.

٧- عاروري، فتحي (٢٠١٣). **المعاينة الإحصائية: طرقها واستخداماتها**. الأردن، عمان: الأكاديميون للنشر والتوزيع.

٨- عبيدات، ذوقان، عدس، عبدالرحمن وعبدالحق، كايد (٢٠٠٢). **البحث العلمي: مفهومه، أدواته، أساليبه**. المملكة العربية السعودية، الرياض: دار أسامة للنشر والتوزيع.

٩- عليان، ربحي مصطفى وغنيم، عثمان محمد (٢٠٠٨). **أساليب البحث العلمي: الأسس النظرية والتطبيق العملي**. الأردن، عمان: دار صفاء للنشر والتوزيع.

١٠- قنديلجي، عامر والسامرائي، إيمان (٢٠٠٩). **البحث العلمي: الكمي والنوعي**. الأردن، عمان: دار اليازوري العلمية للنشر والتوزيع.

١١- زايد، مصطفى (٢٠٠٧). **المرجع الكامل في الإحصاء**. مصر، القاهرة: مطابع الدار الهندسية.

١٢- فهمي، محمد شامل بهاء الدين (٢٠٠٥). **الإحصاء بلا معاناة: المفاهيم مع التطبيقات باستخدام برامج SPSS**. المملكة العربية السعودية، الرياض: معهد الإدارة العامة.

١٣- مراد، صلاح وهادي، فوزية (٢٠١٢). **طرائق البحث العلمي: تصميمها وإجراءاتها**. القاهرة، مصر: دار الكتاب الحديث.

١٤- هندي، محمود وسلمان، خلف (١٤٢٥). **مفاهيم لطرق التحليل الإحصائي**. المملكة العربية السعودية، الرياض: مكتبة الرشد.

ثانياً - المراجع الأجنبية:

1- Agresti, A. & Finlay, B. (2009). ***Statistical methods for the social sciences*** (4th ed.). USA, NJ: Pearson Education, Inc.

2- Antonius, R.(2013). ***Interpreting quantitative data with IBM SPSS Statistics***, (2nd ed.). London: SAGE Publications Ltd.

3- Bankier, M. (1988). Power allocations: determining sample sizes for subnational areas. ***The American Statistician***, 42, 174-177.

4- Bartlett, J. E.; Kotrlik, J. W. & Higgins, C. C. (2001). Organizational research: determining appropriate sample size in survey research. ***Informational Technology, Learning, and Performance Journal***, 19 (1), 44-50.

5- Blueman, A. G. (2011). Elementary statistics: a step by step approach (8th ed.). USA, NY: McGraw Hill.

6- Cattell, R. B. (1966). The scree test for the number of factors. ***Multivariate Behavioral Research***, 1, 629-637.

7- Chuan, C. L.(2006). Sample size estimation using Krejcie and Morgan and Cohen statistical power analysis: a comparison. ***Journal Penyelidikan IPBL***, 7, 79-86.

8- Cohen, J. (1988). ***Statistical power analysis for the behavioral sciences*** (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

9- Cohen, J. (1992). Quantitative methods in psychology: A power primer. ***Psychological Bulletin***, 112 (1), 155–159.

10- Conover, W. J. (1999). ***Practical nonparametric statistics***. USA: John Wiley and Sons.

11- Comrey, A. L., & Lee, H. B. (1992). ***A first course in factor analysis*** (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

12- Dale, A (2006). Quality issues with survey research. ***International Journal of Social Research Methodology***, 9(2), 143-158.

13- Daniel, J.(2012). ***Sampling essentials: practical guidelines for making sampling choices***. USA, CA: Sage Publications.

14- DiStefano, C, Zhu, M & Mindrila, D (2009). Understanding and Using Factor Scores: Considerations for the Applied Researcher. ***Practical Assessment, Research & Evaluation***, 14 (20). Available online: http: //pareonline.net/getvn.asp?v=14&n=20.

15- Field, A. (2005). ***Discovering statistics using SPSS***. London: Sage Publications Ltd.

16- Gay, L. R., & Airasian, P. (2000). ***Educational research: Competencies for analysis and experience*** (6th ed.). Prentice-Hall: USA, N. J.

17- Glass, G. V. & Hopkins, K. D. (1995). ***Statistical methods in education and psychology***. MA: Allyn&Bacon.

18- Gorard, S. (2006). Towards a judgement-based statistical analysis. ***British Journal of Sociology of Education***, 27(1), 67-80.

19- Hair, J. F., Anderson, R.E., and Black, W.C.(1998). ***Multivariate data Analysis*** (5th ed.). Prentice-Hall: USA, N. J.

20- Ho, R. (2006). ***Handbook of univariate and multivariate data analysis and interpretation with SPSS***. USA, Florida: Chapman & Hall/CRC.

21- Huck, S. W.(2012). ***Reading statistics and research*** (6th ed.). Boston: Pearson Education Inc.

22- IBM Corp. Released 2011. ***IBM SPSS Statistics for Windows***, Version 20.0. Armonk, NY: IBM Corp.

23- Johnson, B.& Christensen, L. (2012). ***Educational research: uantitative, qualitative, and mixed approaches*** (4th.). London: Sage Publications Inc.

24- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. ***Educational and Psychological Measurement***, 20, 141-151.

25- Kish, L. (1992). Weighting for unequal pi. ***Journal of Official Statistics***, 8, 183-200.

26- Lajer, K. (2007). Statistical tests as inappropriate tools for data analysis performed on non-random samples of plant communities. ***Folia Geobotanica***, 42, 115-122.

27- Larose, D., (2006). ***Data mining methods and models***. John Wiley & Sons Inc: USA, NJ.

28- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. ***Educational and Psychological Measurement***, 30, 607-610.

29- Madsen, R.(2011). ***Statistics for non-statisticians***. Springer: USA, N. Y.

30- Miles, j. & Shevlin, M. (2007). ***Applying regression & correlation: a guide for students and resrearchers***. London: Sage Publications Inc.

31- Montgomery, D.C.(2012). ***Design and analysis of experiments*** (8th ed.). USA, NY: Whiley.

32- Montgomery, D.C, Peck, E.A. & Vining, G.G.(2006). ***Introduction to linear regression analysis*** (4th ed.). USA, NJ: John Wiley & Sons, Inc.

33- Moore, D. S & Notz, M. I. (2006). ***Statistics: concepts and controversies*** (6th ed.). New York: W. H. Freeman and Company.

34- Salkind, N. J. (2004). ***Statistics for people who (think they) hate statistics*** (2nd ed.). USA, CA: Sage Publications.

35- Snedecor, ***G. W.*** & Cochran, W. G. (1989). S- s (8th ed.). Ames, Iowa: Blackwell Publishing Professional.

36- Spradely, J. P. (1979). ***The Ethnography interview***. CA, USA: Wadsworth.

37- Stevens, S. (1946). On the theory of scales of measurement. ***Science***, 103, 677-680.

38- Suhr, D. (2012). Exploratory Factor Analysis with the World Values Survey. ***Proceedings of the SAS Global Forumm 2012***. Florida.

39- Tabachnick, B. G. & Fidell, L. S. (2007). ***Using multivariate statistics*** (5th ed). Boston: Pearson Education Inc.

40- Thompson, B., (2004). ***Exploratory and confirmatory factor analysis: understanding concepts and applications***. American Psychological Association: Washington, DC.

41- Triola, M. F. (2012). ***Elementary statistics*** (12th ed). Boston: Pearson Education Inc.

42- Warner, R. M. (2008). ***Applied statistics: from bivariate through multivariate techniques***. California: Sage Publication, Inc.

ثالثاً - المواقع الإلكترونية:

1- Anonymous. ***I just Confused Myself Help Please?*** Retrieved Feb 28,2014 from ***http: //www.chegg.com/homework-help/questions-and-answers/us-bureau-labor-andstatistics-reported-person-ages-of18-34-average-92-jobs-see-thisaverage-q1380158***.

2- Anonymous. ***Profits on Sales and Assets***. Retrieved Feb 28,2014 from ***www.uic.edu/classes/idsc/ids472/p1-4.xls***.

3- Anonymous. ***Multiple Regression***. Retrieved Feb 28, 2014 from ***https: //statistics.laerd.com/premium/account.php***.

4- Oskar Blakstad (Mar 10, 2008). ***Research Methodology***. Retrieved Jul 30, 2012 from Explorable.com: ***http: //explorable.com/research-methodology***.

5- Statistics Canada (2010). ***Survey methods and practices***. Retrieved June 20, 2013. from ***http: //www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf***.

الملاحق:

ملحق (١)

دليلك لاختيار الاختبار الإحصائي المناسب

الهدف	نوع المتغير التابع Dependent Variable		
	(فتري Interval أونسبي Ratio وله توزيع طبيعي)	(فتري Interval أونسبي Ratio وله توزيع غيرطبيعي) أو رتبي Ordinal	نوعي ثنائي Dichotomy
مقارنة مجموعة واحدة بقيمة افتراضية	اختبار t لعينة واحدة (one – sample t test)	اختبار ويلكوكسن لعينة واحدة (Wilcoxon signed – rank test)	اختبار مربع كاي (Chi – Square) أو ذات الحدين Binomial
مقارنة مجموعتين مرتبطتين	اختبار t لعينتين مرتبطتين (matched – pair t test)	اختبار ويلكوكسن لعينتين مرتبطتين (Wilcoxon signed – rank test)	اختبار ماكنيمار (McNemar's test)
مقارنة مجموعتين مستقلتين	اختبار t لعينتين مستقلتين (two-sample t test)	اختبار مان – ويتني (Mann – Whitney test)	اختبار مربع كاي (Chi – square) أو اختبار فيشر للعينات الصغيرة (Fisher's test)
مقارنة ٣ مجموعات مستقلة أو أكثر	تحليل التباين في اتجاه واحد (1 – way ANOVA)	اختبار كروسكال – واليس (Krskal – Wallis test)	اختبار مربع كاي (Chi – Square)
مقارنة ٣ مجموعات مرتبطة أو أكثر	تحليل التباين للقياسات المتكررة (repeated – measuresANOVA)	اختبار فريدمان (Friedman test)	اختبار كوكران (Cochran's test)

الهدف	Dependent Variable نوع المتغير التابع		
	(فتري Interval) أونسبي Ratio وله توزيع طبيعي)	(فتري Interval) أونسبي Ratio وله توزيع غيرطبيعي) أو رتبي Ordinal	نوعي ثنائي Dichotomy
القياس الكمي للعلاقة بين متغيرين	ارتباط بيرسون (Pearson Correlation)	ارتباط سبيرمان (Spearman Correlation)	معاملات التوافق (Contingency coefficients)
التنبؤ بقيمة متغير (تابع) بالاعتماد على قيم متغير آخر (مستقل)	الانحدار الخطي أو الغير خطي البسيط (simple linear or non - linear regression)	الانحدار اللامعلمي (nonparametric regression)	الانحدار اللوجستي البسيط (simple logistic regression)
التنبؤ بقيمة متغير (تابع) بالاعتماد على قيم عدة متغيرات أخرى (مستقلة)	الانحدار الخطي أو الغير خطي المتعدد (multiple linear or nonlinear regression)		الانحدار اللوجستي المتعدد (multiple logistic regression)
تخفيض عدد البيانات واستخلاص العوامل	التحليل العاملي Factor Analysis		

ملحق (٢)

أسباب عدم الاستجابة في المسوحات البحثية

عزيزي المتدرب

السلام عليكم ورحمة الله وبركاته،،،

يلعب البحث العلمي دوراً مهماً في التنمية الاقتصادية والاجتماعية. ويعد الاستبيان الأداة الأساسية لجمع البيانات والمعلومات التي تستخدم في البحوث والدراسات الاجتماعية. ويتكون الاستبيان من مجموعة من الأسئلة حول موضوع البحث يُرسل إلى جميع أو بعض أفراد مجتمع البحث. ومع تزايد استخدام الاستبيان كأداة لجمع البيانات من قبل الطلاب والباحثين والشركات في المملكة برزت مشكلة عدم الاستجابة الكلية أو الجزئية للمشاركة في المسوحات البحثية. لذا يهدف هذا البحث إلى معرفة أسباب وعوامل عدم الاستجابة في المسوحات البحثية. وعدم الاستجابة يعني رفض المشاركة تماماً في المسح (أي عدم تعبئة الاستبيان) أو تعبئة جزء من الاستبيان بالإجابة عن بعض الأسئلة.

الاستبيان الذي بين يديك تم تصميمه لمعرفة أسباب عدم الاستجابة من واقع تجربتك كأحد المختارين للمشاركة في المسوحات البحثية في السنوات السابقة. ولتحقيق أهداف البحث يُرجى الإجابة بحرية تامة، إذ تستخدم هذه البيانات لأغراض البحث العلمي فقط. كما نفيدكم بأن تعبئة الاستبيان تستغرق نحو خمس دقائق فقط. شاكرين حسن تعاونكم وتجاوبكم.

ولكم خالص الشكر والتقدير

من خلال خبرتك في السـنوات الماضية، ما مدى اسـتجابتك لتعبئة الاسـتبيانات التي ترد إليك أو إجراء المقابلة معك؟

□ لا أشارك أبداً.

□ أشارك أحياناً.

□ أشارك دائماً.

في حالة المشاركة في تعبئة الاستبيانات التي ترد إليك أو إجراء المقابلة معك؟

□ أقوم بالإجابة عن بعض الأسئلة.

□ أقوم بالإجابة عن معظم الأسئلة.

□ أقوم بالإجابة عن جميع الأسئلة.

إذا كنت لا تشارك دائماً أو تشارك أحياناً في تعبئة الاستبينات، متى تقرر عدم المشاركة؟

□ منذ البداية.

□ بعد الاطلاع على الاستبيان.

يُرجـى تحديد درجة موافقتك على أسـباب عدم الاسـتجابة الكليـة أو الجزئية بالإشارة إلى الإجابة المناسبة بعلامة ().

العبارة	موافق بشدة	موافق	موافق إلى حد ما	غير موافق بشدة	غير موافق
١- الاستبيان طويل من حيث عدد الأسئلة المطلوب الإجابة عنها.					
٢- عدم وضوح الإرشادات اللازمة لتعبئة الاستبيان.					
٣- عدم وضوح الهدف من البحث.					
٤- أسئلة الاستبيان غير متسلسلة (ترتيبها غير منطقي).					
٥- يحتوي الاستبيان على أسئلة مفتوحة كثيرة.					
٦- أسئلة الاستبيان طويلة.					

العبارة	موافق بشدة	موافق	موافق إلى حد ما	غير موافق بشدة	غير موافق
٧- يحتوي الاستبيان على مصطلحات غير مفهومة					
٨- يحتوي الاستبيان على كلمات غير مفهومة					
٩- أسئلة الاستبيان تحتاج إلى تفكير عميق للإجابة عنها					
١٠- يستغرق الاستبيان وقتاً طويلاً للإجابة عن جميع أسئلته.					
١١- يحتوي الاستبيان على أسئلة إيحائية (أسئلة تدل على الإجابة).					
١٢- يحتوي الاستبيان على أسئلة غير ضرورية.					
١٣- يحتوي الاستبيان على أسئلة افتراضية (مثل لو كنت وزيراً ...).					
١٤- الصياغة اللغوية لأسئلة الاستبيان غير سليمة.					
١٥- خيارات الإجابات المحتملة للأسئلة غير كافية.					
١٦- يحتوي الاستبيان على أسئلة ذات خصوصية.					
١٧- يحتوي الاستبيان على أسئلة مركبة (السؤال يتضمن أكثر من سؤال).					
١٨- يحتوي الاستبيان على أسئلة متكررة.					
١٩- ليس لدي الوقت الكافي لتعبئة الاستبيان.					
٢٠- عدم قناعتي بأهداف البحث موضوع الاستبيان.					
٢١- عدم ملاءمة الوقت لتعبئة الاستبيان.					
٢٢- عدم ملاءمة المكان لتعبئة الاستبيان.					

العبارة	موافق بشدة	موافق	موافق إلى حد ما	غير موافق بشدة	غير موافق
٢٣- لا أرى أي أهمية أو مردود للبحث العلمي عموماً.					
٢٤- لدي فكرة سالبة نحو موضوع البحث.					
٢٥- لدي فكرة سالبة نحو المنظمة التي ينتمي إليها الباحث.					
١٦- أخشى من استخدام البيانات التي أدلى بها لغير أغراض البحث.					
٢٧- تتطلب الإجابة عن أسئلة الاستبيان خلفية كبيرة بالموضوع.					
٢٨- أرى أن اختياري غير مناسب للإجابة عن أسئلة الاستبيان.					
٢٩- تصميم الاستبيان غير ملائم.					
٣٠- من سياسة جهة العمل عدم المشاركة في أي استبيان.					

العمر: سنة.

عدد سنوات الخبرة العملية: سنة.

أعلى مؤهل أكاديمي:

☐ أقل من الثانوي.

☐ ثانوي / دبلوم فوق الثانوي.

☐ بكالوريوس.

☐ دبلوم فوق الجامعي / ماجستير.

☐ دكتوراه.

الحالة الاجتماعية:

☐ متزوج.

☐ غير متزوج.

الدخل الشهري بالريال فضلاً حدد: ( ).

ملحق (٣)

المراحل الرئيسية في تحليل بيانات البحوث الكيفية

إن تحليـل البيانات في البحوث الكيفية مشـابهة لتلك التي تسـتخدم في البحوث الكميـة من حيث الإطار العام. فبعد جمع البيانـات وتوثيقها يتم معالجتها وتنظيمها وتصنيفهـا وتحليلها والوصول إلى الاسـتنتاجات والتوصيات. وفـي البحوث الكيفية تحديداً تمر عملية تحليل البيانات بالمراحل التالية (Johnson & Christensen, 2012) و(قنديلجي والسامرائي، ٢٠٠٩):

١- مرحلة تنظيم البيانات Data Organization:

إن قـوة وجودة تحليل البيانات تعتمد على معرفة الباحث جيداً ببيانات دراسـته. ولضمان عدم التشت وفقدان التركيز لابد للباحث من تنظيم الكم الكبير من البيانات التي تم جمعها، وتبويبها بطريقة تسـهل عليه الوصول والرجوع إليها بسـهولة ويسـر واستعداداً لاستخدامها في المراحل اللاحقة من التحليل. ويتم تنظيم البيانات بطرق متعددة حسـب البحث وحسـب ما يراه الباحث مناسباً، فقد يتم تبويب البيانات وفقاً للحالات (مثل الأفراد أو المؤسسـات أو غير ذلك) التي قام بدراستها أو وفقاً لوسائل جمع البيانات المستخدمة مثل الملاحظة، الوثائق، الصور ... إلخ. ويمكن أن يتم تبويب وترتيب البيانات والمعلومات بالطرق اليدوية التقليدية مثل الملفات وبطاقات تسـجيل البيانات والمعلومات أو عن طريق استخدام البرامج الحاسوبية.

٢- مرحلة تصنيف البيانات Categorization وترميزها Coding:

وتهـدف هذه المرحلة إلى تقليص حجـم البيانات إلى الدرجة التـي يجعلها قابلة للإدارة والمعالجة لتحقيق فهـم الظاهرة محل البحث. فبعد تنظيم البيانات وترتيبها يقـوم الباحث بقـراءة بياناته جيداً وبتركيز عالٍ وأكثر من مـرة كي يأخذ فكرة عامة وشاملة عن البيانات، وكي تبرز له نقاط ومفاهيمٌ معينة. بعد ذلك يبدأ بعملية تصنيف أو تجميـع أو تجزيء البيانات في مجموعات أو فئـات Categories معينة قد تتكون مـن كلمـة واحدة أو عبارة أو فقرة، ومن ثم يتم الرمز لهـا برموز معبرة أو وصفها أو تسـميتها. وتتم عملية الترميز بأن يقرأ المحلل أو الباحث البيانات سطراً سطراً ومن ثـم متى وجد الباحث أو المحلل جزءاً من البيانات ذا معنى أو يعبر عن معلومة معينة، يقوم بوصفها أو تسـميتها وإعطائها رمزاً معبراً. وعلى الباحث أن يتذكر دائماً أثناء

عملية التصنيف والترميز الهدف الذي جمعت من أجله البيانات وما هو الشـيء الذي يريد أو يتوقع استنتاجه أو الوصول إليه، وتحديد مجموعة من الأسئلة الرئيسية للإجابة عنها. والجدير بالذكر أن عملية الترميز شاقة وتحتاج مجهود كبير وتركيز ولكنها تعتبر القاعدة الأساسية لعملية التحليل.

وللترميز أو التصنيف عدة أنواع ومن أكثرها استخداماً ما يلي:

١- الرموز المسبقة A Priori Codes: وهي الرموز التي تم إنشاؤها أو بناؤها قبل تحليل البيانات الحالية.

٢- الرموز الاستقرائية Inductive Codes: وهي الرموز التي يقوم الباحث ببنائها أو إنشائها أثناء تحليل البيانات الحالية. وتضم ثلاثة أنواع هي:

- الترميز المفتوح Open Coding: وفيه يتم قراءة النصوص سطراً سطراً ومن ثم تحديد المجموعات والمفاهيم.
- الترميز المحوري Axial Coding: وفيه يتم تنظيم وتصنيف المفاهيم والتصنيفات التي تم تطويرها في الترميز المفتوح وجعلها في تصنيفات أكبر ومفاهيم أكثر تجريداً.
- الترميز الانتقائي Selective Coding: وفيه يتم التركيز على الأفكار والمفاهيم الرئيسية التي تفي بتساؤلات وأهداف الدراسة.

٣- مرحلة تدوين الملاحظات Momoing:

مما لا شك فيه أنه بعد الانتقال من مرحلة التنظيم والوصول إلى مجموعات وفئات تم تكوينها واستخلاصها من البيانات الخام ستكون الصورة للباحث أوضح وستبدأ لدى الباحث بعض الملاحظات والتساؤلات حول تلك المجموعات والفئات التي تم استخلاصها وطبيعة العلاقة بينها، لذا يقوم الباحث بتسجيل تلك الملاحظات والتساؤلات التي يستفيد منها الباحث وتوجهه إلى عمل مزيد من البحث والتدقيق في تلك البيانات والمعلومات المتوافرة وربما الحاجة إلى الحصول على بيانات أو معلومات إضافية تساعده في فهم الظاهرة بشكل أفضل. وعملية تسجيل الملاحظات مستمرة ومتكررة تستمر مع الباحث حتى نهاية البحث، لأنه في كل مرة يقرأ فيها البيانات والمعلومات ويدقق فيها يفهم الباحث موضوع بحثه بشكل أكبر، وهذا يؤدي إلى الحصول على ملاحظات وتساؤلات قيمة ومعلومات جديدة تساعد في تفسير الظاهرة وفهمها بدرجة أفضل.

٤- مرحلة تحديد الأنماط والأفكار الرئيسية Identifying Patterns and Themes والعلاقة بينها:

بعد أن يتم استخلاص المجموعات والفئات في الخطوات السابقة وتدوين الملاحظات والتساؤلات حولها وفهمها واستيعابها بشكل أفضل تبدأ عملية الربط بين تلك المجموعات والفئات وتصنيفها على مستوى أعلى في مجموعات أكبر تسمى بالمفاهيم الرئيسية أو التصنيفات المحورية Axial Coding. وهكذا يمكن أن تستمر العملية في شكل هرمي حتى الوصول إلى المفهوم الأعم والأشمل. ومن هنا نلاحظ أنه في التحليل النوعي - وعلى خلاف التحليل الكمي - يتم التركيز على الروابط والعلاقات بين الفئات والمفاهيم الرئيسية والفرعية التي تم تطويرها، حيث يقوم الباحث بالترميز (أي بناء الفئات والمجموعات Categories) ثم بعد ذلك يبحث عن المفاهيم الرئيسية Themes والأنماط Patterns والعلاقات الارتباطية Realtionships بحيث يعمل على تسجيل الملاحظات Memos والتساؤلات في نفس الوقت ويستمر في ذلك حتى نهاية التحليل.

إن هناك العديد من الأنواع المختلفة لطبيعة العلاقة بين المجموعات والفئات والمفاهيم الرئيسية التي يتم استخلاصها من البيانات الخام. والجدول التالي يوضح أنواع العلاقات الممكنة التي يمكن للباحث أن يتطلع لها والمحتمل وجودها بين الفئات X وY التي تم بناؤها من البيانات الحالية في الدراسة (Spradley, 1979, 111).

اسم العلاقة	شكل العلاقة
الاحتواء الفعلي Strict Inclusion	X تمثل نوع من Y
مكاني Spatial	X هي جزء من Y
سبب وأثر Cause - Effect	X هي سبب في Y أو Y هي نتيجة لـ X
منطقي Rationale	X هي سبب لفعل Y
موقع للإجراء Location for Action	X هي مكان لإجراء Y
وظيفة Function	X تستخدم لإجراء Y
وسيلة Means	X هي طريقة لإجراء Y
تعاقب Sequence	X هي خطوة أو مرحلة لإجراء Y
خاصية Attribution	X هي خاصية أو صفة لـ Y

ومن الوســـائل الأخرى لتحليل العلاقة بين الفئات استخدام المخططات البيانية وهي عبارة عن تمثيل بياني لتوضيح العلاقة بين المفاهيم من خلال استخدام صناديق Boxes (أشكال مستطيلة) وأسهم حيث تحتوي الصناديق المفاهيم أو الفئات، والأسهم تبين طبيعة العلاقة. كما يمكن أيضاً استخدام المصفوفات لتمثيل العلاقة بين الفئات والمفاهيم.

٥- مرحلة تفسير وصياغة النتائج:

وفي هذه المرحلة يقوم الباحث بتفسير نتائج التحليل وصياغتها وفيما يلي مجموعة من الإرشادات التي ينبغي أن يستحضرها الباحث أثناء تفسير وصياغة نتائج البحث وهي:

- على الباحث أن يتذكر جيداً أسئلة بحثه للتأكد من الإجابة عنها.
- ينبغي على الباحث أن يتذكر بأن عملية تفســير النتائج ليســت مجرد وصف فقط للمفاهيم والعلاقات التي اســتنتجها وإنما عليــه أن يتحقق من أن ما توصل إليه له معنى ومهم ويجيب عن أسئلة دراسته ويحقق أهدافها.
- البدء بتلخيص النقاط الرئيسية في موضوع بحثه وذلك بصياغة الأنماط والتصنيفات المحورية والمفاهيم الرئيسية في شكل نتائج للدراسة أو البحث بما يتوافق مع أسئلة وأهــداف بحثه، وهذا ما يطلق عليه بعــض الباحثين الترميز أو التصنيف الانتقائي Selective Coding.
- ما الذي تعلمه الباحث من بحثه الحالي؟ وما الدروس الأساسية التي توصل إليها؟ وما الأشياء الجديدة التي طرأت على البحث أو الدراسة؟
- على الباحث تضمين الآراء الأخرى إن وجدت حول كل مفهوم رئيسي Theme وعدم الاقتصار على تلك التي تدعم رأيه وصب في اتجاه ميوله.
- الاســتعانة بالاقتباس من بيانات البحث أو الدراســة لتوضيح بعض النقاط التي أو المفاهيم التي استنتجها الباحث.
- أن النتائج التي توصل إليها الباحث في هذه المرحلة تبقى على شكل افتراضات.

٦- مرحلة التحقق من النتائج Validating Results:

بعــد أن يصل الباحث إلــى النتائج النهائيــة لبحثه عليه التحقــق من صلاحيتها وموثوقيتها قبل اعتمادها للنشر النهائي. فمن الضروري أن تكون نتائج البحث المنشور ذات إضافــة وقيمة علمية وعملية لحقل المعرفة وللقارئ والمســتفيد. لذا ينبغي على الباحث الرجوع إلى بيانات دراسته أو بحثه بالإضافة إلى العودة مرة أخرى ولو بشكل

ســـريع إلى أدبيات الدراســـة، وذلك للتحقق من النتائج ومناقشتها، وإجراء الإضافات والتعديلات اللازمة، وتوضيح رأيه حول ذلك وتدعيمه قدر الإمكان بالأدلة، والتأكد من مدى توافق التصنيفات والأنماط والافتراضات التي توصل إليها مع بيانات دراسته.

البرامج الحاسوبية المستخدمة في التحليل الكيفي للبيانات:

يوجد العديد من البرامج الحاســـوبية المصممة لتحليل البيانات الكيفية والتي وبلا شك فقد سهلت على الباحثين تنفيذ وعمل معظم الإجراءات الخاصة بتنظيم وتحليل البيانات الكيفية، ومن تلك البرامج ما يلي:

١- Nvivo ويمكن الوصول إليه من خلال الموقع **http: //www.qsrinternational.com**.

٢- ATLAS ويمكن الوصول إليه من خلال الموقع **http: //atlasti.de/**.

٣- NUD-IST ويمكن الوصول إليه من خلال الموقع **http: //www.qsrinternational.com**.

وإلى تاريخ تأليف هذا الكتاب وحســـب حد علم المؤلف، فإن هذه البرامج – وبكل أسف – لا تدعم اللغة العربية.

المؤلف في سطور

- د. سعد بن سعيد القحطاني.
- البريد الإلكتروني: alkahtanisas@yahoo.com

المؤهل العلمي:

- دكتوراه في الإحصاء التطبيقي، جامعة كلورادو الشمالية - الولايات المتحدة الأمريكية ٢٠١٠م.

العمل الحالي:

- أستاذ الإحصاء التطبيقي المساعد. مركز قياس الأداء للأجهزة الحكومية. معهد الإدارة العامة بالرياض.

الخبرات العلمية والعملية:

- التدريب والتدريس لعدد من المواد في مجال الإحصاء والرياضيات تشمل: ضبط الجودة إحصائياً، برنامج التحليل الإحصائي SPSS، الإحصاء المعلمي، الإحصاء اللامعلمي، التحليل الإحصائي للمتغيرات المتعددة، تصميم وتحليل التجارب، تحليل السلاسل الزمنية، الأرقام القياسية، تحليل الارتباط والانحدار الخطي، المعاينة الإحصائية، تصميم الاستبانات، رياضيات البيانات المتقطعة.
- الحضور والمشاركة في العديد من المؤتمرات واللقاءات في مجال الإحصاء.
- تصميم وإعداد عدد من الحقائب التدريبية والملفات العلمية لبرامج تدريبية وإعدادية بمعهد الإدارة العامة.
- تحكيم العديد من البحوث العلمية والاستبانات.
- عضو في لجنة تحديد الحصص التدريبية للأجهزة الحكومية المستفيدة من دورات معهد الإدارة العامة التدريبية.
- عضو لجنة تطوير نظام تقييم البرامج الإعدادية بمعهد الإدارة العامة.
- تقديم استشارة لوزارة الداخلية تتعلق بمراجعة الكتاب الإحصائي السنوي والتقرير المرافق له من حيث الإخراج والمضمون والعمل على تطويره.

- تقديم استشـارة لوزارة العدل تتعلق بمراجعة الكتاب الإحصائي السنوي والتقرير المرافق له من حيث الإخراج والمضمون والعمل على تطويره.
- منسـق وعضو فريق استشـاري من معهد الإدارة العامة لدراسة الحاجة إلى إنشاء وحدة لقياس الأداء في المجلس الأعلى للقضاء.
- عضو لجنة البحوث الدائمة بمعهد الإدارة العامة منذ عام ٢٠١١م إلى الآن.
- مستشار إحصائي غير متفرغ بوزارة الداخلية منذ عام ٢٠١١م الى ٢٠١٣م.
- عضـو اللجنة العلمية بمركز أبحـاث الجريمة بوزارة الداخليـة منذ بداية ٢٠١٣م إلى الآن.
- مستشار إحصائي غير متفرغ بوزارة العمل منذ منتصف ٢٠١٣م إلى الآن.
- مدير مشروع قياس رضا المستفيدين من خدمات الأجهزة الحكومية في مركز قياس الأداء للأجهزة الحكومية بمعهد الإدارة العامة.
- مستشار إحصائي لمدة ثلاث سنوات (٢٠٠٧م - ٢٠٠٩م) بمركز الاستشارات للبحوث بجامعة كلورادو الشمالية.

أوراق العمل والأعمال العلمية الأخرى المقدمة:

- تم تقديم ورقة عمل بعنوان «خريطة ضبط المتوسط المتحرك الموزون أسّياً للمتغيرات المتعددة لمراقبة متوسط العمليات» في مؤتمر اللقاء الإحصائي المشترك في واشنطن بالولايات المتحدة الأمريكية عام ٢٠٠٩م.
- تم تقديم ورقة عمل بعنوان «خريطة ضبط المتوسط المتحرك الموزون أسّياً للمتغيرات المتعددة لمراقبة متوسط العمليات» في محاضرة مفتوحة في جامعة كلورادو الشمالية عام ٢٠٠٩م.
- تم تقديم ورقة عمل بعنوان «المفلتر الخطي الأسّـي غير المتحيز للسلاسل الزمنية» في مؤتمر اللقاء الإحصائي المشترك بولاية كلورادو بالولايات المتحدة الأمريكية عام ٢٠٠٨م.
- تم تقديم ورقة عمل بعنوان «المفلتر الخطي الأسّـي غير المتحيز للسلاسل الزمنية» في محاضرة مفتوحة في جامعة كلورادو الشمالية عام ٢٠٠٨ م.

- المؤلف الرئيســـي لورقة علمية منشورة في مجلة Communications in Statistics- Simulation and Computation الأمريكيـــة بعنـــوان "Double multivariate exponentially weighted moving average control chart for process mean monitoring" عام ٢٠١١ م بالاشتراك مع البروفيسور Jay Shaffer.
- ورقة علمية منشـــورة في المجلة الإحصائية المصرية بعنوان "Linear exponential- smoother of time series"، ٢٠١٢م.
- ورقة علمية منشورة في مجلة Journal of Modern Applied Statistical Methods بعنـــوان "Robustness of DEWMA versus EWMA Control Charts to Non- Normal Processes"، عام ٢٠١٣م.

الجوائز والمكافآت:

- مكافأة تميز علمي من معهد الإدارة العامة عام ٢٠١٤م.
- جائـــزة عميـــد الدراســـات العليا للتفوق الأكاديمـــي في مرحلة الدكتـــوراه في عام ٢٠٠٩م.
- الحصـــول علـــى الميدالية الذهبية كجائـــزة لأفضل أطروحة دكتوراه على مســـتوى الجامعة عام ٢٠١٠م.

تصميم وإخراج وطباعة
الإدارة العامة للطباعة والنشر بمعهد الإدارة العامة - ١٤٣٦هـ

هذا الكتاب

يهدف وبشكل أساسي إلى تسليط الضوء على المفاهيم الأساسية وأساليب وأدوات التحليل الإحصائي الأكثر استخداماً في البحوث والدراسات العلمية لاسيما الاجتماعية منها والإنسانية، وتقديمها بطريقة عملية تطبيقية بحتة تمكن الباحثين والمستفيدين غير المتخصصين في علم الإحصاء من استيعاب تلك المفاهيم والأساليب وتطبيقها بسهولة ويسر.

وقد تميز الكتاب عن غيره من الكتب العربية في مجال الإحصاء التطبيقي في عدة جوانب، منها: أن الأساليب الإحصائية الوصفية والاستدلالية التي تضمنها هذا الكتاب تم انتقاؤها بما يوفر احتياج البحوث والدراسات العلمية من أساليب التحليل الإحصائي الأكثر استخداماً وتطبيقاً في الواقع العملي، ومن ثم طرحها بطريقة تطبيقية ميسرة ومباشرة - يمكن للباحثين والمستفيدين استخدامها كنموذج عملي.

كما أن الكتاب اشتمل على العديد من "الإضاءات الإحصائية" التي تسلط الضوء على المفاهيم والأساليب الإحصائية التي يساء استخدامها من قبل الباحثين والطلاب والممارسين والتساؤلات الجدلية حول بعض القضايا الإحصائية، حيث تناول وناقش تحديد الحد الأدنى لأحجام العينات العشوائية البسيطة أو المركبة والعينات غير العشوائية اللازم توافرها لإجراء الدراسات والبحوث التي تحتوي على العديد من المتغيرات الكمية والنوعية المختلفة. كما تم تسليط الضوء على اختبارات الفرضيات في حالة بيانات الحصر الشامل وأنسب اختبارات المقارنات البعدية في تحليل التباين، وغيرها من التساؤلات والمفاهيم الأخرى التي تم مناقشتها وتوضيح الرأي حولها وتضمينها في كتاب واحد يوفر على الباحث عناء التنقل بين الكتب الأجنبية والعربية للحصول على استيضاح حول تلك القضايا والمفاهيم.


* 9 9 6 0 1 4 2 3 4 5 *

تصميم وإخراج وطباعة الإدارة العامة للطباعة والنشر- معهد الإدارة العامة ١٤٣٦هـ